

# Overcoming universal restrictions on metal selectivity by protein design

<https://doi.org/10.1038/s41586-022-04469-8>

Tae Su Choi<sup>1</sup> & F. Akif Tezcan<sup>1,✉</sup>

Received: 17 June 2021

Accepted: 21 January 2022

Published online: 2 March 2022

 Check for updates

Selective metal coordination is central to the functions of metalloproteins:<sup>1,2</sup> each metalloprotein must pair with its cognate metallocofactor to fulfil its biological role.<sup>3</sup> However, achieving metal selectivity solely through a three-dimensional protein structure is a great challenge, because there is a limited set of metal-coordinating amino acid functionalities and proteins are inherently flexible, which impedes steric selection of metals.<sup>3,4</sup> Metal-binding affinities of natural proteins are primarily dictated by the electronic properties of metal ions and follow the Irving–Williams series<sup>5</sup> ( $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$ ) with few exceptions<sup>6,7</sup>. Accordingly, metalloproteins overwhelmingly bind  $Cu^{2+}$  and  $Zn^{2+}$  in isolation, regardless of the nature of their active sites and their cognate metal ions<sup>1,3,8</sup>. This led organisms to evolve complex homeostatic machinery and non-equilibrium strategies to achieve correct metal speciation<sup>1,3,8–10</sup>. Here we report an artificial dimeric protein,  $(AB)_2$ , that thermodynamically overcomes the Irving–Williams restrictions in vitro and in cells, favouring the binding of lower-Irving–Williams transition metals over  $Cu^{2+}$ , the most dominant ion in the Irving–Williams series. Counter to the convention in molecular design of achieving specificity through structural preorganization,  $(AB)_2$  was deliberately designed to be flexible. This flexibility enabled  $(AB)_2$  to adopt mutually exclusive, metal-dependent conformational states, which led to the discovery of structurally coupled coordination sites that disfavour  $Cu^{2+}$  ions by enforcing an unfavourable coordination geometry. Aside from highlighting flexibility as a valuable element in protein design, our results illustrate design principles for constructing selective metal sequestration agents.

Several strategies have been developed for building artificial metalloproteins<sup>11–15</sup>, including the de novo design of new protein folds and complexes<sup>16–18</sup>, repurposing of proteins for incorporation of metal coordination sites<sup>19,20</sup> and metal-templated construction of supramolecular protein assemblies<sup>21,22</sup>. These approaches have yielded many examples of functional metalloproteins that are assembled in vitro (that is, in homogeneous solutions). However, building from scratch a protein that selectively binds a transition metal ion over others in a heterogeneous environment represents a demanding task. Rare successes in this arena are either based on the redesign of the interiors of pre-existing proteins<sup>23,24</sup> or do not exhibit complete counter-Irving–Williams selectivity<sup>25</sup>. This difficulty stems from the fact that de novo protein design approaches have yet to provide the requisite structural precision (well below 0.5 Å) needed to sterically discriminate between transition metal ions, which is exacerbated by the challenge of computationally modelling metal–protein interactions<sup>26,27</sup>. Furthermore, our predictive understanding of how protein structure and metal coordination influence one another is limited, which poses a fundamental problem for any deterministic metalloprotein design campaign with a ‘single structure–single function’ goal.

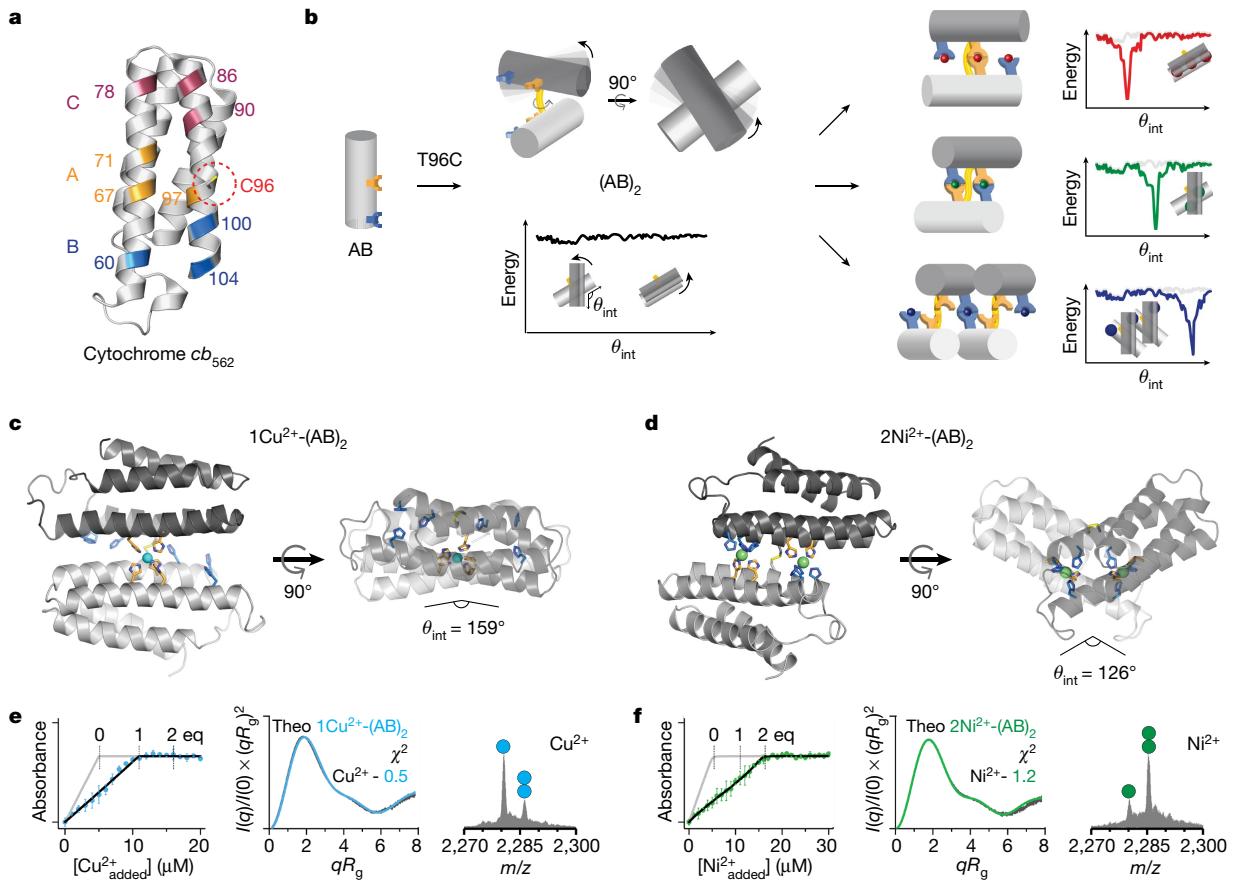
Against this backdrop, we set out to pursue an alternative, probabilistic approach to metalloprotein design, which relies on a flexible

protein–protein interface that is replete with metal-binding residues. Our inspiration came from natural multi-metallic proteins (for example, metallothionein<sup>28</sup>, calmodulins<sup>29</sup> and calprotectin<sup>30</sup>) with inherent flexibilities that enable structural coupling between different metal coordination sites. We were particularly intrigued by the selective  $Mn^{2+}$ -over- $Fe^{2+}$  binding in the dinuclear sites of class Ib ribonucleotide reductases and the suggestion that such counter-Irving–Williams selectivity may be achieved by the cooperativity between the two sites<sup>6,7</sup>. Accordingly, we predicted that structural coupling between multiple metal sites within a flexible protein scaffold could engender enhanced binding affinities or selectivities.

## Design and characterization of $(AB)_2$

We previously introduced a strategy termed MASCoT (metal active sites through covalent tethering)<sup>31</sup>. In MASCoT, a protein–protein interface is simply created by the disulfide-mediated linking of two copies of a monomeric protein, and metal-binding motifs are subsequently incorporated near the disulfide bond to generate interfacial metal coordination sites. The disulfide ‘pivot’ furnishes a malleable assembly that possesses a shallow free-energy landscape and is capable of exploring different conformations. However, the assembly is also structurally

<sup>1</sup>Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA. <sup>✉</sup>e-mail: tezcan@ucsd.edu



**Fig. 1 | Design and characterization of the (AB)<sub>2</sub> scaffold.** **a**, Locations of metal coordination motifs (A, B and C) and C96 on the cytochrome *cb*<sub>562</sub> surface. **b**, Design of disulfide-linked (AB)<sub>2</sub> with a malleable dimer interface, along with cartoon representations of possible metal-free and metal-bound conformations and corresponding free-energy landscapes. **c, d**, Crystal structures of 1Cu<sup>2+</sup>-(AB)<sub>2</sub> (**c**) and 2Ni<sup>2+</sup>-(AB)<sub>2</sub> (**d**). Cu<sup>2+</sup> and Ni<sup>2+</sup> ions are represented as cyan and green spheres, respectively. X-ray data collection and refinement statistics are listed in Extended Data Table 2. **e, f**, Characterization of 1Cu<sup>2+</sup>-(AB)<sub>2</sub> (**e**) and 2Ni<sup>2+</sup>-(AB)<sub>2</sub> (**f**) complexes in solution using competitive

metal titrations, solution SAXS and ESI-MS. Log-scale SAXS plots are shown in Extended Data Fig. 2j–m. Left, changes in Fura-2 absorbance at 335 nm are plotted with theoretical metal-binding isotherms (grey) in the absence of (AB)<sub>2</sub>. Experimental data points and error bars are presented as mean and s.d. of three independent measurements. Middle, theoretical SAXS profiles derived from the crystal structures of 1Cu<sup>2+</sup>-(AB)<sub>2</sub> (cyan) and 2Ni<sup>2+</sup>-(AB)<sub>2</sub> (green) are plotted with experimental SAXS profiles (black) and corresponding  $\chi^2$  values. Right, circles in ESI-MS spectra represent the number of Ni<sup>2+</sup> (green) and Cu<sup>2+</sup> (cyan) ions bound to (AB)<sub>2</sub>.

constrained by the disulfide bond, and, upon metal binding, forms a well-ordered interface. In proof-of-principle experiments, cytochrome *cb*<sub>562</sub> (a monomeric, four-helix bundle haem-protein<sup>32</sup>) was tailored with a single Cys (C96) and a proximal tris-His (3His) motif comprising H67, H71 and H97 (Fig. 1a). The resulting dimer, (Cys3His)<sub>2</sub>, was observed to coordinate all divalent, mid-to-late first row transition metal ions in an identical, distorted square-pyramidal geometry, consisting of five of the available six His residues (H67/H67', H71/H71' and H97). The M<sup>2+</sup>-(Cys3His)<sub>2</sub> dissociation constants ( $K_d$ ) ranged from 20  $\mu\text{M}$  for Mn<sup>2+</sup> to 400 fM for Cu<sup>2+</sup>, following the Irving–Williams series<sup>31</sup>.

Here, to access alternative dimer conformations with multiple metal-binding sites, we engineered three cytochrome *cb*<sub>562</sub> variants (AB, AC and BC) bearing two 3His motifs surrounding the C96 pivot. These variants combine two of the following three 3His sites in addition to the native H63: motif A (H67, H71 and H97), motif B (H60, H100 and H104) and motif C (H78, H86 and H90). These His-enriched variants were isolated in high yields from bacterial cultures and formed stable disulfide-linked dimers upon oxidation. Analytical ultracentrifugation and solution small-angle X-ray scattering (SAXS) experiments indicated that the metal-free AB dimer was monodisperse at all concentrations tested, but AC and BC variants appeared to form some higher-order oligomers or aggregates and were therefore not studied further (Extended Data Fig. 1).

We anticipated that the flexible AB dimer, (AB)<sub>2</sub>, could form several different metal-dependent conformations (Fig. 1b). To investigate these possibilities, we first sought to crystallize (AB)<sub>2</sub> in the presence of first-row transition metal ions (Mn<sup>2+</sup>, Fe<sup>2+</sup>, Co<sup>2+</sup>, Ni<sup>2+</sup>, Cu<sup>2+</sup> and Zn<sup>2+</sup>). Of these, we determined the crystal structures of the Co<sup>2+</sup>-, Ni<sup>2+</sup>- and Cu<sup>2+</sup>-(AB)<sub>2</sub> complexes at high resolution (better than 2.0 Å). The Cu<sup>2+</sup>-(AB)<sub>2</sub> adduct was isostructural with the previously described M<sup>2+</sup>-(Cys3His)<sub>2</sub> complexes (root-mean-square deviation (RMSD) over all C<sub>α</sub> atoms  $\approx 0.5$  Å), featuring a flat, antiparallel dimer interface with an intermonomer angle  $\theta_{\text{int}} = 159^\circ$  (Fig. 1c). We refer to this dimer geometry as conformation-1. In conformation-1, only a single, central 5His coordination site consisting of two A motifs participated in Cu<sup>2+</sup> coordination (electron paramagnetic spectrum shown in Extended Data Fig. 2a), and the two B motifs were unoccupied. By contrast, Co<sup>2+</sup>- and Ni<sup>2+</sup>-(AB)<sub>2</sub> complexes possessed two symmetrical metal coordination sites on either side of the C96–C96 linkage, each composed of A and B motifs in a square-pyramidal 5His geometry, with H60, H100 and H104 from one monomer and H67' and H71' from the opposing monomer (Fig. 1d, Extended Data Fig. 2b). To accommodate this new metal-binding configuration, the monomers in 2Co<sup>2+</sup>- and 2Ni<sup>2+</sup>-(AB)<sub>2</sub> complexes (RMSD = 0.27 Å with respect to one another) rotate by more than 30° about the C96–C96 linkage relative to the Cu<sup>2+</sup>-(AB)<sub>2</sub> structure, yielding a kinked architecture with  $\theta_{\text{int}} = 126^\circ$  (conformation-2).

To examine whether the crystal structures of  $M^{2+}$ - $(AB)_2$  complexes reflected their solution structures, we carried out metal-binding assays and SAXS studies (Fig. 1e, f, Extended Data Fig. 2c, Extended Data Table 1). Titrations using Fura-2 as a competitive indicator confirmed that  $(AB)_2$  tightly bound a single equivalent of  $Cu^{2+}$  and two equivalents of  $Co^{2+}$  or  $Ni^{2+}$ , with fM and nM  $K_d$  values, respectively.  $Mn^{2+}$  and  $Fe^{2+}$  (which lie low in the Irving–Williams series) or  $Zn^{2+}$  (which typically disfavours all-His, high-coordination-number geometries) did not show appreciable binding affinities or selectivities over  $Cu^{2+}$  (Extended Data Fig. 2d–f). Given that the  $(AB)_2$  adducts of these ions also did not yield crystal structures, we did not consider them further. For SAXS analyses, we used the crystal structures of the  $Cu^{2+}$ - and  $2Co^{2+}$ - or  $2Ni^{2+}$ - $(AB)_2$  complexes to calculate theoretical scattering profiles. The calculated SAXS pattern of  $Cu^{2+}$ - $(AB)_2$  (conformation-1) is quite distinct from those of  $2Co^{2+}$ - and  $2Ni^{2+}$ - $(AB)_2$  (conformation-2), as expected from their different crystal structures (Extended Data Fig. 2g). The experimental SAXS profiles for each species ( $Cu^{2+}$  versus  $Ni^{2+}$  or  $Co^{2+}$ ) matched their corresponding theoretical profiles closely ( $\chi^2 = 0.5\text{--}1.8$ ) but deviated from those of the other species ( $\chi^2 = 9.3\text{--}14.1$ ), indicating that the crystallographically observed conformations of  $(AB)_2$  complexes were populated in solution (Extended Data Fig. 2h, i). These observations established that  $(AB)_2$  formed two distinct and mutually exclusive metal-dependent conformations in solution.

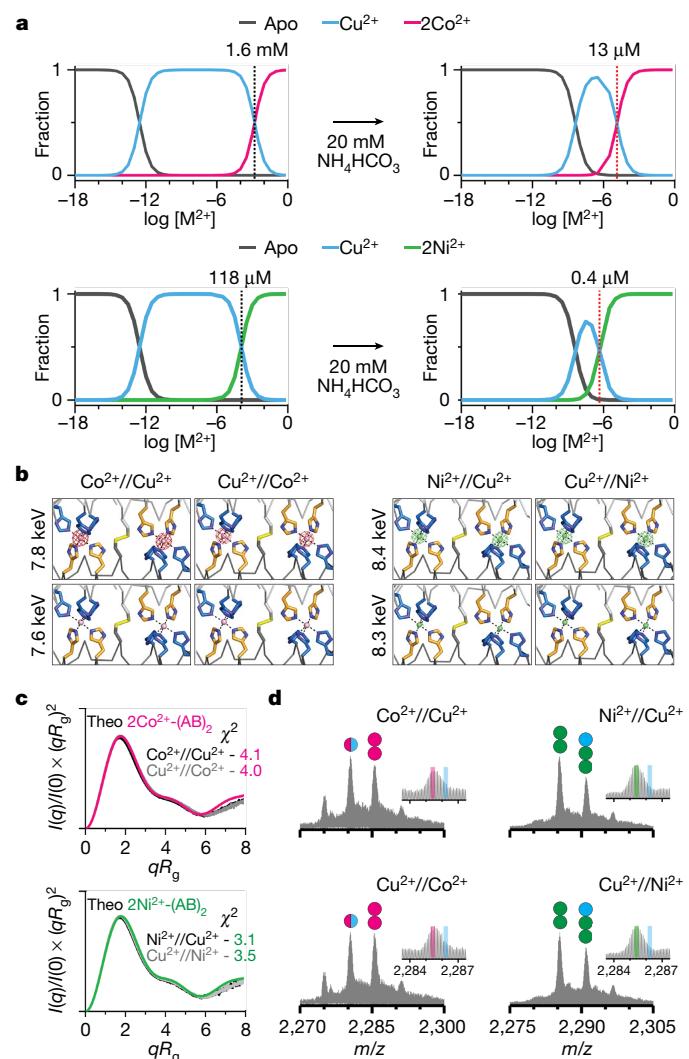
### Counter-Irving–Williams selectivity of $(AB)_2$

Although the per-site  $Cu^{2+}$  binding affinity of  $(AB)_2$  is more than 10,000-fold higher than that for  $Co^{2+}$  and  $Ni^{2+}$ , we predicted that  $Co^{2+}$  and  $Ni^{2+}$  might outcompete  $Cu^{2+}$  given their higher valency and potential cooperativity. Equilibrium calculations based on the experimentally determined  $M^{2+}$ - $(AB)_2$   $K_d$  values indicated that the population of  $2Co^{2+}$ - and  $2Ni^{2+}$ - $(AB)_2$  complexes should indeed begin to dominate the  $Cu^{2+}$ - $(AB)_2$  species at  $[M^{2+}]_{\text{total}} > 10^{-3}\text{--}10^{-4}$  M (Fig. 2a, Supplementary Information). Given that these concentrations are typical of crystallization conditions, we first examined the metal selectivity of  $(AB)_2$  in the crystalline state.  $(AB)_2$  (2 mM) was crystallized in a mixture of 4 mM  $Cu^{2+}$  and 4 mM  $Co^{2+}$  or  $Ni^{2+}$ . In these experiments, the protein solution was incubated first either with  $Cu^{2+}$  or with  $Co^{2+}$  or  $Ni^{2+}$  for at least one hour, after which the second metal ion was added and the resulting mixture was left to equilibrate for one day. Under all four conditions ( $Co^{2+}/Cu^{2+}$ ,  $Cu^{2+}/Co^{2+}$ ,  $Ni^{2+}/Cu^{2+}$  and  $Cu^{2+}/Ni^{2+}$ ), the observed crystal structures were identical to those of  $2Co^{2+}$ - and  $2Ni^{2+}$ - $(AB)_2$  (that is, conformation-2), with two peripheral metal coordination sites (Extended Data Fig. 3a). Anomalous X-ray diffraction experiments confirmed that these sites were populated exclusively by Co or Ni (Fig. 2b).

In parallel, we prepared similar samples for solution SAXS experiments. The Kratky plots of these samples showed close correspondence with the theoretical scattering patterns of  $2Co^{2+}$ - and  $2Ni^{2+}$ - $(AB)_2$  structures in conformation-2 ( $\chi^2 = 3.5\text{--}4.1$ ) (Fig. 2c), but not with that of  $Cu^{2+}$ - $(AB)_2$  in conformation-1 ( $\chi^2 = 17.0\text{--}22.4$ ) (Extended Data Fig. 3b), confirming that  $Co^{2+}$  and  $Ni^{2+}$  fully outcompete  $Cu^{2+}$  in solution under thermodynamic control.

To investigate whether the counter-Irving–Williams selectivity of  $(AB)_2$  was operative under lower protein and metal concentrations, we turned to native electrospray ionization mass spectrometry (ESI–MS)<sup>33</sup>. ESI–MS of 5  $\mu$ M  $(AB)_2$  in 20 mM  $NH_4HCO_3$  (pH 7.8) without metal ions showed a broad charge-state distribution from +10 to +24 (Extended Data Fig. 3c). We chose the most abundant, +11 state for further analyses of  $M^{2+}$ - $(AB)_2$  complexes, whereby two molar equivalents of  $M^{2+}$  (10  $\mu$ M) were added to  $(AB)_2$ . Consistent with crystal structures and competitive titrations, we observed  $1Cu^{2+}$ - $(AB)_2$ ,  $2Co^{2+}$ - $(AB)_2$  and  $2Ni^{2+}$ - $(AB)_2$  as the predominant species, with an abundance of greater than 80% (Fig. 1e, f, Extended Data Fig. 2c).

For ESI–MS under competitive binding conditions, we prepared similar  $Co^{2+}/Cu^{2+}$ ,  $Cu^{2+}/Co^{2+}$ ,  $Ni^{2+}/Cu^{2+}$  and  $Cu^{2+}/Ni^{2+}$  samples, using



**Fig. 2 |  $Co^{2+}$  and  $Ni^{2+}$  selectivity of  $(AB)_2$ .** **a**, Theoretical fractions of metal-free and metal-bound  $(AB)_2$  species calculated from  $K_d$  values and buffered metal concentrations in 20 mM MOPS and 150 mM NaCl (left) and 20 mM  $NH_4HCO_3$  (right). **b**, Anomalous scattering densities of interfacial metal sites of  $(AB)_2$ . Anomalous densities contoured at 5.0  $\sigma$  were collected around the K-edge of Co (around 7.7 keV) and Ni (around 8.3 keV). X-ray data collection and refinement statistics are listed in Extended Data Table 3. **c**, Experimental Kratky plots of  $(AB)_2$  (black or grey) compared with theoretical plots of  $2Co^{2+}$ - $(AB)_2$  (magenta) and  $2Ni^{2+}$ - $(AB)_2$  (green) structures (log-scale plots are shown in Extended Data Fig. 3f–g). **d**, ESI–MS spectra of  $(AB)_2$  under  $Co^{2+}/Cu^{2+}$ ,  $Cu^{2+}/Co^{2+}$ ,  $Ni^{2+}/Cu^{2+}$  and  $Cu^{2+}/Ni^{2+}$  competition conditions. Circles in ESI–MS spectra represent the number of  $Co^{2+}$  (magenta),  $Ni^{2+}$  (green) and  $Cu^{2+}$  (cyan) ions bound to  $(AB)_2$ . Half circles (magenta and cyan) indicate a mixture of  $Co^{2+}$  and  $Cu^{2+}$  complexes. Insets represent expanded  $m/z$  ranges of  $2M^{2+}$ - $(AB)_2$  complexes, with magenta, green and cyan lines corresponding to theoretical  $m/z$  values of  $2Co^{2+}$ ,  $2Ni^{2+}$ , and  $2Cu^{2+}$ - $(AB)_2$  complexes, respectively.

micromolar  $(AB)_2$  and  $M^{2+}$  (Fig. 2d). In  $Co^{2+}/Cu^{2+}$  and  $Cu^{2+}/Co^{2+}$  samples, we observed approximately equal amounts of  $2Co^{2+}$ - $(AB)_2$  and  $1M^{2+}$ - $(AB)_2$  species, with the latter attributed to a mixture of  $1Co^{2+}$ - $(AB)_2$  and  $1Cu^{2+}$ - $(AB)_2$ . These results indicated that  $Co^{2+}$  could still compete with  $Cu^{2+}$  for binding  $(AB)_2$  at  $\mu$ M concentrations. Notably, in the  $Ni^{2+}/Cu^{2+}$  and  $Cu^{2+}/Ni^{2+}$  samples,  $2Ni^{2+}$ - $(AB)_2$  was the dominant species, with no evidence for a  $1Cu^{2+}$ - $(AB)_2$  complex. The analysis of a sample containing 5  $\mu$ M  $(AB)_2$  and substoichiometric amounts of  $Cu^{2+}$  (5  $\mu$ M) and  $Ni^{2+}$  or  $Co^{2+}$  (5  $\mu$ M) revealed roughly equal populations of  $1Cu^{2+}$ - $(AB)_2$  and  $2Ni^{2+}$ - $(AB)_2$  or  $2Co^{2+}$ - $(AB)_2$  complexes, with no indication of heterobimetallic

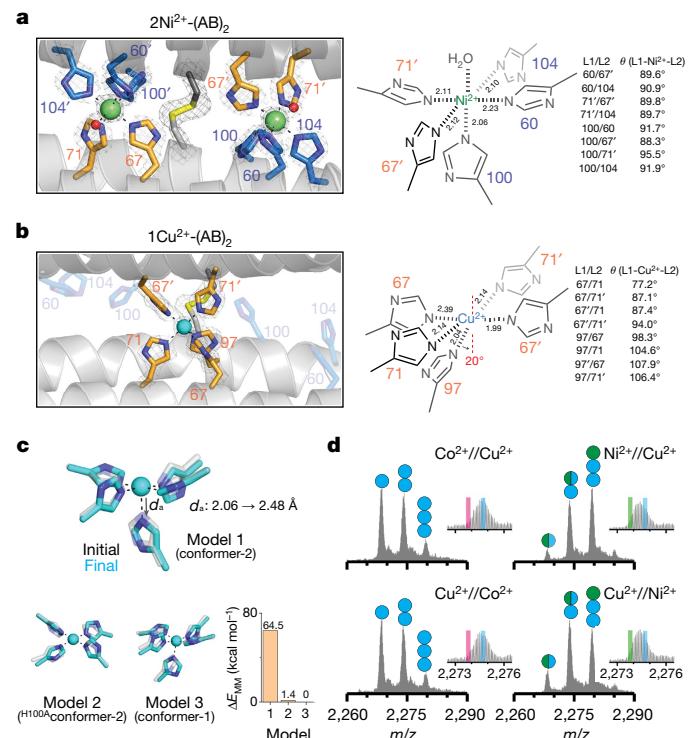
species (Extended Data Fig. 3d). This observation supports our structural findings that two equivalents of  $\text{Ni}^{2+}$  or  $\text{Co}^{2+}$  cooperatively bind to the peripheral sites of conformation-2 in an all-or-nothing manner, whereas  $\text{Cu}^{2+}$  only coordinates conformation-1 in one equivalent.

The equilibrium fractionation profiles shown in Fig. 2a suggested that  $\text{Ni}^{2+}$  and  $\text{Co}^{2+}$  should not be able to outcompete  $\text{Cu}^{2+}$  for  $(\text{AB})_2$  binding at  $\mu\text{M}$  concentrations. However, we considered that the ESI–MS samples also contained 20 mM  $\text{NH}_4\text{HCO}_3$  as a volatile buffer and the observed  $\text{M}^{2+}\text{-}(\text{AB})_2$  populations must also be influenced by the metal– $\text{NH}_3$  binding equilibria: complexation with  $\text{NH}_3$  would be expected to lower the concentration of free  $\text{Cu}^{2+}$  in solution to a greater extent compared to  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$ , owing to the considerably higher formation constants of  $\text{Cu}^{2+}$ -amine complexes<sup>34</sup>. Indeed, competitive binding titrations of  $(\text{AB})_2$  performed in 20 mM  $\text{NH}_4\text{HCO}_3$  showed that the apparent  $K_d$  for the  $\text{Cu}^{2+}\text{-}(\text{AB})_2$  complex decreased by around 15,000-fold, whereas those for the  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  complexes decreased by less than 10-fold (Extended Data Fig. 3e, Extended Data Table 1). The revised equilibrium fractionation profiles (calculated using the apparent  $K_d$  values in 20 mM  $\text{NH}_4\text{HCO}_3$ ) show that  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  begin dominating  $\text{Cu}^{2+}$  binding at around 13  $\mu\text{M}$  and around 0.4  $\mu\text{M}$  concentrations, respectively, consistent with ESI–MS results (Fig. 2a). These results illustrate that competing equilibria that involve exchange-labile ligands in solution can amplify the metal selectivity of a protein scaffold. This strategy is similar to what is observed in cells, which contain a surplus of competing ligands that buffer intracellular metals to the inverse of the Irving–Williams series<sup>35,36</sup>.

## Basis for counter-Irving–Williams selectivity

Having established  $(\text{AB})_2$  as a unique protein scaffold that defies the Irving–Williams series, we next sought to understand the structural basis for its selectivity. It should be noted at the outset that the  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  selectivity cannot arise from their higher valency of binding compared to  $\text{Cu}^{2+}$ , because the two peripheral, 5His coordination motifs in conformation-2 are also available in solution for two  $\text{Cu}^{2+}$  ions to bind. The fact that: (a)  $\text{Cu}^{2+}$  binds singly to the central 5His site in conformation-1 and (b)  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  exclusively occupy the peripheral 5His sites over  $\text{Cu}^{2+}$  in conformation-2 mandates that each peripheral site must individually disfavour  $\text{Cu}^{2+}$  binding (that is, each site binds  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  more strongly than  $\text{Cu}^{2+}$ ). Because both the central and peripheral sites are 5His motifs, we surmised that any selectivity must arise from steric effects, which prompted us to closely inspect the coordination geometries. The peripheral sites possess an almost-perfect octahedral geometry, with nearly equidistant  $\text{M}^{2+}$ –ligand bonds ( $2.17 \pm 0.06 \text{ \AA}$  for  $\text{Co}^{2+}$  and  $2.12 \pm 0.06 \text{ \AA}$  for  $\text{Ni}^{2+}$ ) and minimal deviation from ideal octahedral bond angles ( $90.5 \pm 2.0^\circ$  and  $90.9 \pm 2.2^\circ$ ) (Fig. 3a, Extended Data Fig. 4a). All His side chains are positioned for optimal bonding between their  $N_e$  atoms and the metal, including the axial H100 ligand. By contrast, the geometry of the central site is distorted from an octahedron with a broad distribution of bond lengths ( $2.16 \pm 0.16 \text{ \AA}$  for  $\text{Cu}^{2+}$ ) and angles ( $95.3 \pm 11.9^\circ$ ). Notably, the imidazole group of the axial H97 ligand is slanted by nearly  $20^\circ$  with respect to the equatorial plane normal (Fig. 3b), presenting a less than ideal  $\sigma$ -donation geometry to  $\text{Cu}^{2+}$ .

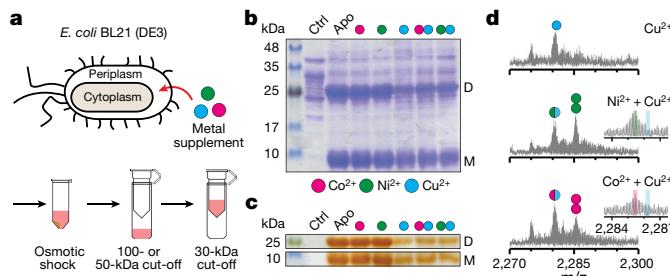
Thus, we hypothesized that the peripheral sites in conformation-2 might disfavour  $\text{Cu}^{2+}$  binding by rigidly enforcing an ideal octahedral coordination geometry (featuring the  $\sigma$ -donating axial H100 ligand), thereby preventing the energetically favoured, tetragonal distortion of the  $\text{Cu}^{2+}$  centre. To further examine this hypothesis, we first analysed the protein environment around H100 and found that the imidazole side chain had no alternative rotameric configurations that could be adopted without sterically clashing with nearby side chains and protein backbone (Extended Data Fig. 4b–d). Of note, any tetragonal distortion in one peripheral site would necessitate a conformational change in the entire dimeric assembly, thereby deforming the second



**Fig. 3 | Primary metal coordination spheres of central and peripheral sites in  $(\text{AB})_2$ .** **a, b**, Magnified views of 5His coordination of  $\text{Ni}^{2+}$  (a) and  $\text{Cu}^{2+}$  (b) with the  $2\text{mF}_c\text{-DF}_c$  electron density map (grey mesh) contoured at  $5.0\sigma$  (metal) or  $1.5\sigma$  (ligand). Coordination distances and angles between  $\text{M}^{2+}$  and ligands are shown on the right. **c**, QM-MM optimized  $\text{Cu}^{2+}$  coordination in model 1 ( $\text{Cu}^{2+}$  in the peripheral site of  $(\text{AB})_2$ ), model 2 ( $\text{Cu}^{2+}$  in the peripheral site of  ${}^{\text{H}100\text{A}}(\text{AB})_2$ ) and model 3 ( $\text{Cu}^{2+}$  in the central site of  $(\text{AB})_2$ ). The initial (crystal structures) and final (QM-MM optimized structures) conformations of models 1–3 are coloured grey and cyan, respectively.  $d_s$  represents the bond distance between  $\text{Cu}^{2+}$  and the  $\text{N}_e$  atom of H100. Relative MM energies of  $(\text{AB})_2$  conformations resulting from the optimized models 1–3 are presented in the bar graph. **d**, ESI–MS spectra of  ${}^{\text{H}100\text{A}}(\text{AB})_2$  under competitive metal-binding conditions. Circles in ESI–MS spectra represent the number of  $\text{Ni}^{2+}$  (green) and  $\text{Cu}^{2+}$  (cyan) ions bound to  $(\text{AB})_2$ . Half circles (green and cyan) indicate the mixture of  $\text{Ni}^{2+}$  and  $\text{Cu}^{2+}$  complexes. Insets show expanded  $m/z$  ranges for  $2\text{M}^{2+}\text{-}(\text{AB})_2$  complexes with magenta, green and cyan lines corresponding to theoretical  $m/z$  values of  $2\text{Co}^{2+}$ ,  $2\text{Ni}^{2+}$  and  $2\text{Cu}^{2+}\text{-}{}^{\text{H}100\text{A}}(\text{AB})_2$  complexes, respectively.

peripheral site. This expectation was borne out by quantum mechanics and molecular mechanics (QM/MM) calculations (Fig. 3c), in which we considered three structural models of  $(\text{AB})_2$  as initial states for simulations: model 1: conformation-2 with a single  $\text{Cu}^{2+}$  ion substituted into one of the peripheral metal-binding sites based on the crystal structure of  $2\text{Ni}^{2+}\text{-}(\text{AB})_2$ ; model 2: the same as model 1 but with a H100A mutation to eliminate axial coordination; model 3: the crystal structure of  $\text{Cu}^{2+}\text{-}(\text{AB})_2$  (conformation-1). In the case of model 1, QM (B3LYP) and MM (AMBER) optimization led to a marked distortion of  $\text{Cu}^{2+}$  coordination, with an increase in the axial H100– $\text{Cu}^{2+}$  distance from 2.06 to 2.48  $\text{\AA}$ . This distortion was accompanied by a large conformational change in the dimer structure, eliminating the other peripheral binding site and leading to an increase in the overall energy of the system (Fig. 3c, Extended Data Fig. 5). By contrast, models 2 and 3 did not undergo significant changes in the  $\text{Cu}^{2+}$  coordination geometries and overall  $(\text{AB})_2$  conformations, and both were considerably more stable compared to model 1 (Extended Data Fig. 5).

Next, we experimentally evaluated the effects of the H100A mutation. In a reversal of the metal-binding properties of  $(\text{AB})_2$ , ESI–MS experiments indicated that the  ${}^{\text{H}100\text{A}}(\text{AB})_2$  dimer bound two equivalents



**Fig. 4 | Selective Co<sup>2+</sup> and Ni<sup>2+</sup> binding of (AB)<sub>2</sub> in the *E. coli* periplasmic space.** **a**, Extraction and purification scheme of periplasmic AB species.

**b, c**, SDS-PAGE of periplasmic extracts stained with Coomassie blue for all proteinaceous contents (**b**) and *o*-dianisidine for haem-proteins (**c**); supplemented metal ions are represented as magenta (Co<sup>2+</sup>), green (Ni<sup>2+</sup>) and cyan (Cu<sup>2+</sup>). D and M refer to the (AB)<sub>2</sub> dimer and the AB monomer, respectively. Reproducibility of SDS-PAGE was tested using three independently extracted sample sets. **d**, ESI-MS spectra of (AB)<sub>2</sub> extracted from periplasm of the cells grown in Cu<sup>2+</sup>, Co<sup>2+</sup> + Cu<sup>2+</sup> or Ni<sup>2+</sup> + Cu<sup>2+</sup> conditions. Circles in ESI-MS spectra represent the number of Co<sup>2+</sup> (magenta), Ni<sup>2+</sup> (green) and Cu<sup>2+</sup> (cyan) ions bound to (AB)<sub>2</sub>. Half circles (green and cyan, and magenta and cyan) indicate the mixture of metal-bound (AB)<sub>2</sub>. Insets show expanded *m/z* ranges for 2M<sup>2+</sup>-(AB)<sub>2</sub> complexes with magenta, green and cyan lines corresponding to theoretical *m/z* values of 2Co<sup>2+</sup>, 2Ni<sup>2+</sup> and 2Cu<sup>2+</sup>-(AB)<sub>2</sub> complexes, respectively.

of Cu<sup>2+</sup>, but only one equivalent of Co<sup>2+</sup> or Ni<sup>2+</sup> (Extended Data Fig. 6a). Fura-2 competition assays further indicated that H100A(AB)<sub>2</sub> coordinated two equivalents of Cu<sup>2+</sup> (Extended Data Fig. 6b). Crystal structures of the Co<sup>2+</sup>- and Ni<sup>2+</sup>-H100A-(AB)<sub>2</sub> complexes revealed that they adopted conformation-1 with a single metal ion in the central 5His site (Extended Data Fig. 6c, d). Competition experiments showed that the H100A mutation completely abolished Co<sup>2+</sup> and Ni<sup>2+</sup> selectivity over Cu<sup>2+</sup>, with the latter ion dominating binding in all conditions tested (Fig. 3d). Collectively, our results confirmed the counterintuitive mechanism by which (AB)<sub>2</sub> destabilizes Cu<sup>2+</sup> binding through an increased coordination number and enforced octahedral geometry, thereby achieving counter-Irving–Williams selectivity.

### In vivo metal selectivity of (AB)<sub>2</sub>

Selective incorporation of metals in the cellular environment is essential for the development of artificial metalloproteins and metalloenzymes with in vivo activities<sup>21,37</sup>. With such synthetic biological applications in mind, we investigated whether (AB)<sub>2</sub> showed metal selectivity in the periplasm of bacterial cells. AB is equipped with an N-terminal leader sequence for translocation to the periplasm, in which the oxidizing environment would allow the formation of the disulfide-mediated (AB)<sub>2</sub> dimer<sup>21</sup>. We expressed AB in *Escherichia coli* (BL21) cells, and then incubated the intact cells in a growth medium supplemented with 20 μM of Co<sup>2+</sup>, Ni<sup>2+</sup>, Cu<sup>2+</sup> or a mixture Co<sup>2+</sup> + Cu<sup>2+</sup> and Ni<sup>2+</sup> + Cu<sup>2+</sup> (Fig. 4a). The cells were collected and subjected to osmotic shock for the periplasmic extraction of protein components, which were analysed for metal content. Gel electrophoretic analyses indicated that AB exists as a mixture of roughly equal amounts of dimers and monomers (Fig. 4b, c) and is trapped in the periplasm without secretion to the external medium (Extended Data Fig. 7a, b). ESI-MS measurements of the extracts from Co<sup>2+</sup>- and Ni<sup>2+</sup>-supplemented cells indicated both singly and doubly M<sup>2+</sup>-bound (AB)<sub>2</sub> complexes (Extended Data Fig. 7c), whereas the (AB)<sub>2</sub> complexes that were extracted from Cu<sup>2+</sup>-supplemented cells contained one Cu<sup>2+</sup> equivalent (Fig. 4d), paralleling the findings of in vitro studies. Under competitive conditions with a mixture of Ni<sup>2+</sup> + Cu<sup>2+</sup>, 2Ni<sup>2+</sup>-(AB)<sub>2</sub> was detected as the major species (around 60%) with a smaller population (around 40%) of singly M<sup>2+</sup>-bound (AB)<sub>2</sub> complex (Fig. 4d). The population of the doubly bound, 2Co<sup>2+</sup>-(AB)<sub>2</sub> species

was reduced (around 35%) compared to singly bound species (around 65%), probably owing to the lower Co<sup>2+</sup> affinity of (AB)<sub>2</sub>. Consistent with a metal-exchange equilibrium between the periplasm and the external medium, when the metal concentrations in the lysogeny broth (LB) medium were raised to 100 μM in competition conditions, the periplasmic fractions of 2Co<sup>2+</sup>- and 2Ni<sup>2+</sup>-(AB)<sub>2</sub> increased to more than 80% (Extended Data Fig. 7d). These observations show that the in vitro, metal selectivity of (AB)<sub>2</sub> for Co<sup>2+</sup> and Ni<sup>2+</sup> over Cu<sup>2+</sup> is also operative in the heterogeneous environment of the periplasm and controllable through metal concentrations in the external milieu.

### Conclusions

Metal selectivity by natural proteins relies on a precise arrangement of amino acids that is frequently associated with preorganized architectures. For example, the rigid selectivity filter in the potassium ion channel can distinguish between K<sup>+</sup> and Na<sup>+</sup> ions, the ionic radii of which differ by only 0.36 Å<sup>38</sup>. Yet there are also prominent examples such as calmodulin<sup>29</sup> or metalloregulatory<sup>39</sup> proteins that are inherently flexible but become well-structured after binding their cognate metals to reach the same level of sub-Å precision. These features are realized in the design of (AB)<sub>2</sub>, the ‘acquired rigidity’ of which is driven by direct structural coupling between symmetrical coordination sites in de novo interfaces. (AB)<sub>2</sub> was not designed with the specific goal of selecting against Cu<sup>2+</sup> binding at the outset. However, the disulfide-bridged (AB)<sub>2</sub> interface was deliberately constructed to be flexible, yet conducive to yielding structurally well-ordered states, which was essential for the discovery of a new design principle for selective metal coordination. This mode of construction is reminiscent of natural protein folds that simultaneously possess structural order and flexibility to enable the emergence of new functions<sup>40</sup> and cooperativity or allostery<sup>41,42</sup>. We believe that such flexible, minimally tethered and evolutionarily unconstrained protein–protein interfaces provide a suitable platform for the design and directed evolution of selective binding sites for metal ions as well as other substrates.

Our findings also show that it is possible to thermodynamically overcome the near-universal dominance of the Irving–Williams series solely through the design of a three-dimensional protein architecture. However, achieving selectivity over Cu<sup>2+</sup> comes at the cost of a rigidified protein structure and a highly restrictive metal coordination geometry, which leaves little room for developing higher-order metal-based functions such as catalysis, electron transfer and signalling that rely on diverse metal coordination environments. By extension, this implies that the bioinorganic complexity of a cell could not have evolved solely by relying on the three-dimensional structures of proteins to selectively bind metal ions under thermodynamic control, therefore necessitating actively driven metal homeostasis mechanisms for the emergence of complex organisms.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-022-04469-8>.

1. Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. Metalloproteins and metal sensing. *Nature* **460**, 823–830 (2009).
2. Gray, H. B., Stiefel, E. I., Valentine, J. S. & Bertini, I. *Biological Inorganic Chemistry: Structure and Reactivity* (University Science Books, 2007).
3. Waldron, K. J. & Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **7**, 25–35 (2009).
4. Dudev, T. & Lim, C. Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. *Chem. Rev.* **114**, 538–556 (2014).
5. Frausto da Silva, J. J. R. & Williams, R. J. P. *The Biological Chemistry of the Elements* (Oxford University Press, 2001).

6. Kisgeropoulos, E. C. et al. Key structural motifs balance metal binding and oxidative reactivity in a heterobimetallic Mn/Fe protein. *J. Am. Chem. Soc.* **142**, 5338–5354 (2020).

7. Gråve, K., Griese, J. J., Berggren, G., Bennett, M. D. & Högbom, M. The *Bacillus anthracis* class Ib ribonucleotide reductase subunit NrdF intrinsically selects manganese over iron. *J. Biol. Inorg. Chem.* **25**, 571–582 (2020).

8. Reyes-Caballero, H., Campanello, G. C. & Giedroc, D. P. Metalloregulatory proteins: metal selectivity and allosteric switching. *Biophys. Chem.* **156**, 103–114 (2011).

9. O'Halloran, T. V. & Culotta, V. C. Metallochaperones, an intracellular shuttle service for metal ions. *J. Biol. Chem.* **275**, 25057–25060 (2000).

10. Tottey, S. et al. Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature* **455**, 1138–1142 (2008).

11. Lombardi, A., Pirro, F., Maglio, O., Chino, M. & DeGrado, W. F. De novo design of four-helix bundle metalloproteins: one scaffold, diverse reactivities. *Acc. Chem. Res.* **52**, 1148–1159 (2019).

12. Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855–862 (2009).

13. Yu, F. et al. Protein design: toward functional metalloenzymes. *Chem. Rev.* **114**, 3495–3578 (2014).

14. Schwizer, F. et al. Artificial metalloenzymes: reaction scope and optimization strategies. *Chem. Rev.* **118**, 142–231 (2018).

15. Churchfield, L. A. & Tezcan, F. A. Design and construction of functional supramolecular metalloprotein assemblies. *Acc. Chem. Res.* **52**, 345–355 (2019).

16. Faiella, M. et al. An artificial di-iron oxo-protein with phenol oxidase activity. *Nat. Chem. Biol.* **5**, 882–884 (2009).

17. Zastrow, M. L., Peacock, F. A., Stuckey, J. A. & Pecoraro, V. L. Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.* **4**, 118–123 (2012).

18. Studer, S. et al. Evolution of a highly active and enantiospecific metalloenzyme from short peptides. *Science* **362**, 1285–1288 (2018).

19. Khare, S. D. et al. Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat. Chem. Biol.* **8**, 294–300 (2012).

20. Yeung, N. et al. Rational design of a structural and functional nitric oxide reductase. *Nature* **462**, 1079–1082 (2009).

21. Song, W. J. & Tezcan, F. A. A designed supramolecular protein assembly with in vivo enzymatic activity. *Science* **346**, 1525–1528 (2014).

22. Churchfield, L. A., Medina-Morales, A., Brodin, J. D., Perez, A. & Tezcan, F. A. De novo design of an allosteric metalloprotein assembly with strained disulfide bonds. *J. Am. Chem. Soc.* **138**, 13163–13166 (2016).

23. Zhou, L. et al. A protein engineered to bind uranyl selectively and with femtomolar affinity. *Nat. Chem.* **6**, 236–241 (2014).

24. Wegner, S. V., Boyaci, H., Chen, H., Jensen, M. P. & He, C. Engineering a uranyl-specific binding protein from NikR. *Angew. Chem. Int. Ed. Engl.* **48**, 2339–2341 (2009).

25. Brodin, J. D. et al. Evolution of metal selectivity in templated protein interfaces. *J. Am. Chem. Soc.* **132**, 8610–8617 (2010).

26. Guffy, S. L., Der, B. S. & Kuhlman, B. Probing the minimal determinants of zinc binding with computational protein design. *Protein Eng. Des. Sel.* **29**, 327–338 (2016).

27. Akcapinar, G. B. & Sezerman, O. U. Computational approaches for de novo design and redesign of metal-binding sites on proteins. *Biosci. Rep.* **37**, BSR20160179 (2017).

28. Byrd, J. & Winge, D. R. Cooperative cluster formation in metallothionein. *Arch. Biochem. Biophys.* **250**, 233–237 (1986).

29. Halling, D. B., Liebeskind, B. J., Hall, A. W. & Aldrich, R. W. Conserved properties of individual  $\text{Ca}^{2+}$ -binding sites in calmodulin. *Proc. Natl Acad. Sci. USA* **113**, E1216–E1225 (2016).

30. Zygiel, E. M. & Nolan, E. M. Transition metal sequestration by the host-defense protein calprotectin. *Annu. Rev. Biochem.* **87**, 621–643 (2018).

31. Rittle, J., Field, M. J., Green, M. T. & Tezcan, F. A. An efficient, step-economical strategy for the design of functional metalloproteins. *Nat. Chem.* **11**, 434–441 (2019).

32. Faraone-Mennella, J., Tezcan, F. A., Gray, H. B. & Winkler, J. R. Stability and folding kinetics of structurally characterized cytochrome c- $b_{562}$ . *Biochemistry* **45**, 10504–10511 (2006).

33. Choi, T. S., Lee, H. J., Han, J. Y., Lim, M. H. & Kim, H. I. Molecular insights into human serum albumin as a receptor of amyloid- $\beta$  in the extracellular region. *J. Am. Chem. Soc.* **139**, 15437–15445 (2017).

34. Burgot, J.-L. Ionic Equilibria in Analytical Chemistry (Springer, 2012).

35. Osman, D. et al. Bacterial sensors define intracellular free energies for correct enzyme metallation. *Nat. Chem. Biol.* **15**, 241–249 (2019).

36. Young, T. R. et al. Calculating metallation in cells reveals CobW acquires  $\text{Co}^{II}$  for vitamin  $\text{B}_{12}$  biosynthesis while related proteins prefer  $\text{Zn}^{II}$ . *Nat. Commun.* **12**, 1195 (2021).

37. Jeschek, M. et al. Directed evolution of artificial metalloenzymes for *in vivo* metathesis. *Nature* **537**, 661–665 (2016).

38. Thompson, A. N. et al. Mechanism of potassium-channel selectivity revealed by  $\text{Na}^+$  and  $\text{Li}^+$  binding sites within the KcsA pore. *Nat. Struct. Mol. Biol.* **16**, 1317–1324 (2009).

39. Capdevila, D. A., Braymer, J. J., Edmonds, K. A., Wu, H. & Giedroc, D. P. Entropy redistribution controls allostery in a metalloregulatory protein. *Proc. Natl Acad. Sci. USA* **114**, 4424–4429 (2017).

40. Tokuriki, N. & Tawfik, D. S. Protein dynamism and evolvability. *Science* **324**, 203–207 (2009).

41. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allostery. *Nature* **508**, 331–339 (2014).

42. Papaleo, E. et al. The role of protein loops and linkers in conformational dynamics and allostery. *Chem. Rev.* **116**, 6391–6423 (2016).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

## Methods

### Protein mutagenesis, expression and purification

A pET20b(+)/Cys(His)<sub>3</sub> vector with ampicillin resistance was used for site-directed mutagenesis of the original Cys(His)<sub>3</sub> variant into (AB)<sub>2</sub> variants<sup>31</sup>. PCR products generated from site-directed mutagenesis were transformed into chemically competent XL-1 blue cells (Agilent) and plated onto LB-agar plates with 100 mg l<sup>-1</sup> ampicillin at 37 °C for 24 h. Plasmids were purified using Express plasmid Miniprep kit (Biomiga) and sequenced (Eton Bioscience). Sequenced plasmids were transformed into chemically competent BL21(DE3) cells containing a *ccm* (cytochrome c maturation) cassette with chloramphenicol resistance and plated onto LB-agar plates containing 100 mg l<sup>-1</sup> ampicillin and 33 mg l<sup>-1</sup> chloramphenicol at 37 °C for 24 h<sup>32,43</sup>. Isolated colonies were transferred to a 5 ml LB medium with 100 mg l<sup>-1</sup> ampicillin and 33 mg l<sup>-1</sup> chloramphenicol and grown for 12–14 h. These cultures were inoculated into 2.8-l glass flasks containing 1.3 l LB medium with 100 mg l<sup>-1</sup> ampicillin and 33 mg l<sup>-1</sup> chloramphenicol. Flasks were shaken at 100 rpm for 16–24 h at 37 °C. Pink-coloured cells were collected using centrifugation at 6,000g for 7 min, and the cell pellet was stored at -80 °C. Frozen cell pellets were suspended in a 10 mM sodium acetate buffer solution (50–100 ml, pH 4.5) with vigorous stirring for 1 h. The pellets were lysed by repeated cycles of sonication in an ice bath. The pH of the lysate was adjusted to 10 and lowered to 4.5 to precipitate cell components other than cyt cb<sub>562</sub> variants. After centrifugation (10,000 rpm for 10 min), red-coloured supernatant was filtered using a 0.45-μm syringe disk-filter. The clear supernatant was diluted with a 10 mM sodium acetate buffer solution (0.8 l, pH 4.5) and applied to a CM Sepharose Fast Flow resin pre-equilibrated with the same buffer. The column was washed with additional buffer (0.2 l) and eluted using a step-gradient from 0 to 1,000 mM NaCl. Fractions with  $A_{415}/A_{280}$  values > 4 were pooled, concentrated and exchanged into a 10 mM sodium acetate buffer solution (pH 4.5). The protein was then purified on a High-S cartridge column using a step-gradient of 0 to 1,000 mM NaCl using a Biologic DuoFlow workstation (Bio-Rad). Fractions with  $A_{415}/A_{280}$  values > 6 were pooled, concentrated and exchanged into a 20 mM MOPS buffer solution (pH 7.4) using a Vivaspin-6 3-kDa centrifugal filter (Sartorius). CuCl<sub>2</sub> (0.5× protein monomer concentration) was added to catalyse the formation of disulfide bond. The mixture of protein and CuCl<sub>2</sub> was incubated at 37 °C for 8 h. The protein was treated with 10-fold excess EDTA overnight at 4 °C to prepare metal-free protein. EDTA and monomeric protein were removed using a Chelex100-treated, 20 mM MOPS buffer solution (pH 7.4) in a Vivaspin-6 10-kDa centrifugal filter (Sartorius). The purity of the protein was confirmed by SDS-PAGE and  $A_{415}/A_{280}$  values<sup>44</sup>. Concentrated proteins (around 10 mM) were kept at 4 °C and used within a month.

### SAXS

SAXS scattering profiles of (AB)<sub>2</sub> were obtained at Beamline 4-2 of the Stanford Synchrotron Radiation Laboratory (SSRL). The sample-to-detector distance was set to 1.7 m, and the temperature was maintained at 20 °C. Protein concentrations were set to 20 mg ml<sup>-1</sup> (800 μM dimer) in a 20 mM MOPS buffer solution (pH 7.4) containing 150 mM NaCl. All samples were transferred to 8×PCR strip tubes and placed in the BL4-2 autosampler<sup>45</sup>. The autosampler injected 25 μl of the solution into a quartz capillary tube with a diameter of 1.5 mm equipped in the lab-made sample holder. SAXS profiles of the samples were recorded five times (1 pattern per s) with 2 s intervals to monitor the structural transition induced by the incident X-ray. No X-ray induced changes were observed in the samples of (AB)<sub>2</sub> variants. Then, SAXS scattering patterns of two independently prepared protein solutions were collected using Blu-Ice 5 software. All measurements were averaged and background-subtracted in SAXSPipe. Experimental SAXS profiles were reproducible under identical conditions after 6 months.

Further data analysis of SAXS profiles was performed using the ATSAS software package<sup>46</sup>. Intensities of all merged scattering profiles were normalized in SAS data analysis. Guinier analysis of (AB)<sub>2</sub>, (AC)<sub>2</sub>, and (BC)<sub>2</sub> variants was performed using the equation below in SAS data analysis software of ATSAS package<sup>46</sup>.

$$\ln[I(q)] = \ln[I(0)] - \frac{1}{3}q^2R_g^2,$$

where  $q$  and  $I(q)$  are the scattering vector and the scattering intensity. Theoretical scattering profiles of 2Co<sup>2+</sup>-(AB)<sub>2</sub>, 2Ni<sup>2+</sup>-(AB)<sub>2</sub> and 1Cu<sup>2+</sup>-(AB)<sub>2</sub> structures were generated in CRYSTAL<sup>47</sup> using default parameters and were compared with experimental scattering profiles. Chi-square ( $\chi^2$ ) values were evaluated from differences between experimental scattering profiles and theoretical scattering profiles using the equation below:

$$\chi^2 = \frac{1}{m-1} \sum_{k=1}^m \left( \frac{I_{\text{theo}}(q) - I_{\text{exp}}(q)}{\sigma_{\text{exp}}(q)} \right)^2,$$

where  $m$  is the total number of points in scattering profiles, and  $\sigma_{\text{exp}}$  is the standard deviation of experimental scattering profiles.  $\sigma_{\text{exp}}$  values were obtained using repeat exposures of the same sample. Experimental scattering profiles were converted to normalized Kratky plots using the function below to emphasize a  $q$ -range from 0.1 to 0.3 Å<sup>-1</sup>, which corresponds to the change of dimer orientation:

$$(qR_g)^2 I(q)/I(0) \text{ versus } qR_g$$

### X-ray crystallography

Red-coloured crystals of (AB)<sub>2</sub> variants were obtained by sitting-drop vapour diffusion at room temperature. Crystallization conditions are listed in Supplementary Table 1. Protein solutions were preincubated with the metal ion of interest for 1 h. For crystallization under metal competition conditions, protein solutions were incubated with the first metal ion for 1 h. Then the second metal ion was added and the solution was incubated for an additional 24 h. Sitting drop contained 1–2 μl each of the protein and precipitant solutions and the reservoir contained 500 μl of the precipitant solution. Crystals typically appeared within a week and were collected after 2–8 weeks. Experimental crystal growth conditions are provided in Supplementary Table 1. Concentrations of dimer and metal ions were 2 mM and 4 mM, respectively. Perfluoro polyether (Hampton) was used for cryoprotection before crystal freezing in liquid N<sub>2</sub>. Diffraction data were collected at ALS Beamline 5.0.2, and SSRL Beamline 9-2 using multi-wavelength radiation (Blu-Ice 5 software) or at the UCSD crystallography facility using a Bruker APEX II Cu-edge source diffractometer at 100 K (APEX3 software). Diffraction patterns of three to five protein crystals in one to two screening conditions were confirmed without significant changes in overall protein structures, and the highest-resolution dataset was chosen for further analysis. Diffraction data were processed using iMosflm and Scala<sup>48</sup>. Molecular replacement was performed using Phaser-MR with monomeric cyt cb<sub>562</sub> (Protein Data Bank (PDB): 2BC5) as a search model. Model refinement using phenix.refine and metal/water placement using Coot<sup>49</sup> were performed to generate final models. Electron density maps were generated using PHENIX<sup>50</sup> and all graphics related to structures and electron densities were produced using PyMOL<sup>51</sup>. X-ray data collection and refinement statistics are listed in Extended Data Tables 2 and 3.

### Analytical ultracentrifugation

To confirm the oligomeric state of initially designed (AB)<sub>2</sub>, (AC)<sub>2</sub>, and (BC)<sub>2</sub> variants, analytical ultracentrifugation (AUC) analysis was performed using 25 μM protein dimer in a buffer solution containing

20 mM MOPS and 150 mM NaCl (pH 7.4). Sedimentation velocities were measured on a Beckman XL-A instrument using an An-60 Ti rotor. For each sample, 250–600 scans were obtained at 41,000 rpm and 25 °C for 12–16 h using ProteoLab software. All data were processed using SEDFIT software<sup>52</sup> with the following fixed parameters:  $d_{\text{buffer}} = 1.0049 \text{ g ml}^{-1}$ ;  $\eta_{\text{buffer}} = 0.010214 \text{ poise}$ ;  $V_{\text{bar}} = 0.73084$ .

### ESI-MS

ESI-MS experiments were performed using an Elite hybrid linear ion trap-orbitrap mass spectrometer equipped with a heated-electrospray ionization (HESI) source (Thermo Fisher Scientific). All solutions for ESI-MS samples were prepared using LC-MS grade water (JT Baker). The concentrations of (AB)<sub>2</sub> dimers were adjusted to 5 μM in 20 mM NH<sub>4</sub>HCO<sub>3</sub>. Protein solutions were exchanged using 20 mM to remove non-specifically bound metal ions. The capillary voltage was adjusted to 2.5 kV, the flow rate was set to 3–10 μl min<sup>-1</sup>, the source temperature was set to 100 °C and the gas flow rates (sheath and aux) were set to 12. For each sample, 200 spectra were obtained for 3–4 min and averaged for further analysis using Xcalibur software. To prevent cross-contamination of metal ions, peek tubing and steel capillary were flushed out using 0.1% formic acid and 20 mM NH<sub>4</sub>HCO<sub>3</sub> in each measurement. High mass mode at 240,000 resolution was applied to obtain the best resolution limit. Estimated resolution ( $m/\Delta m$ ) of the mass spectrometer was approximately 80,000 around +11 charge state of (AB)<sub>2</sub>. To assure reproducibility of ESI-MS results, two independent experiments were performed with the interval of one to two weeks. Theoretical and experimental *m/z* values of (AB)<sub>2</sub> and <sup>110</sup>OA(AB)<sub>2</sub> complexes are listed in Supplementary Tables 2–5.

### Periplasmic extraction of (AB)<sub>2</sub>

BL21(DE3) colonies with pET20B(+) vector encoding (AB)<sub>2</sub> were inoculated into a 25 ml LB medium in 50-ml Falcon tubes and shaken at 200 rpm for 24 h at 37 °C. Then, metal ions were added to the LB medium, which was incubated at 200 rpm for additional 8 h at 37 °C. Cells were centrifuged at 8,000g for 10 min at 4 °C. The periplasmic extraction protocol was adopted from previous reports<sup>21,53</sup> with small modifications. The cell pellets were washed with 4 ml ice-cold 1× phosphate buffered saline (PBS). No Co<sup>2+</sup>, Ni<sup>2+</sup>, and Cu<sup>2+</sup> contaminant was observed in the ICP-MS analysis of the fresh 1× PBS solution. Only negligible amounts (less than 10 nM) of Co<sup>2+</sup>, Ni<sup>2+</sup> and Cu<sup>2+</sup> were detected in the 1× PBS washing step, indicating that nearly all non-intracellular metal ions were removed by the cell pelleting and the washing step. The pellet was resuspended in 1.5 ml of a 20 mM NH<sub>4</sub>HCO<sub>3</sub> solution with 20% (w/v) sucrose and transferred to 2-ml tubes. The suspension was incubated for 20 min at room temperature and centrifuged at 17,000g for 1 min. The pellet was then resuspended with 0.75 ml of ice-cold 20 mM NH<sub>4</sub>HCO<sub>3</sub> solution and incubated for 20 min at room temperature to induce the lysis of the outer membrane, followed by centrifugation at 17,000g for 2 min. Freshly prepared solutions of 20 mM NH<sub>4</sub>HCO<sub>3</sub> (with or without 20% sucrose) were also free of Co<sup>2+</sup>, Ni<sup>2+</sup> and Cu<sup>2+</sup> contaminants, as determined by inductively coupled plasma mass spectrometry (ICP-MS) analysis. The red-coloured supernatant was diluted to 2 ml using 20 mM NH<sub>4</sub>HCO<sub>3</sub> and filtered using a Vivaspin-6 100-kDa centrifugal filter (Sartorius) to remove cell debris and large-molecular-weight components. The extract was filtered using an Amicon ultra-0.5 50-kDa centrifugal filter (Millipore). The filtrate was concentrated with an Amicon ultra-0.5 30-kDa centrifugal filter (Millipore) and buffer-exchanged using a 20 mM NH<sub>4</sub>HCO<sub>3</sub> solution 4–6 times to remove low-molecular-weight species. The  $A_{415}/A_{280}$  ratio was in the range of 2–4. ESI-MS spectra of purified extracts were obtained under the same parameters as above. ESI-MS experiments of the extracts were performed in duplicate using two independent samples to examine reproducibility of metal selectivity in the periplasm.

### QM/MM calculations

To evaluate the conformational stabilities of Cu<sup>2+</sup>-bound (AB)<sub>2</sub> structures, QM/MM calculations of Cu<sup>2+</sup>-bound (AB)<sub>2</sub> were performed using two-layer ONIOM (our own *n*-layered integrated molecular orbital and molecular mechanics) approach<sup>54,55</sup>. Using Chimera 1.13.1, water boxes with 1 nm × 1 nm × 1 nm dimensions were created around the structures of model 1 (Cu<sup>2+</sup> in the peripheral site of (AB)<sub>2</sub>), model 2 (Cu<sup>2+</sup> in the peripheral site of <sup>110</sup>OA(AB)<sub>2</sub>) and model 3 (Cu<sup>2+</sup> in the central site of (AB)<sub>2</sub>). These structures were converted to Gaussian input files using the TAO package<sup>56</sup>. 5His/4His metal coordination spheres were designated as the QM region and the remaining parts of the model structures were set as the MM region. QM calculations were performed using the B3LYP<sup>57,58</sup> function with 6-31G+ (d,p) basis set<sup>59–61</sup> and MM calculations were performed using the AMBER<sup>62,63</sup> force field equipped in Gaussian 09. The interactions between QM and MM parts were treated using mechanical embedding<sup>64,65</sup>. Force field for the haem group is not defined in AMBER, but even without the haem group, the cyt cb<sub>562</sub> scaffold was highly stable in a pilot test of MM optimization. Thus, all input files were prepared without the haem group. The QM region was released during optimization to find the most stable Cu<sup>2+</sup>-bound 5His/4His coordination sites whereby the MM region was optimized. After QM-MM optimization of models 1–3, water molecules were removed and single-point energy calculation was performed to evaluate relative stabilities of the optimized structures. The MM energy ( $E_{\text{MM}}$  (model)) of the QM (model) part was subtracted from the MM energy ( $E_{\text{MM}}$  (real)) of the whole (real) system to correct the energy of the MM region. Because the total number of atoms in the MM region of models 1–3 was identical,  $E_{\text{MM}}$  (real) –  $E_{\text{MM}}$  (model) values were used to compare conformational stabilities of (AB)<sub>2</sub> in models 1–3. Model 3 showed the lowest energy value in the MM region, so model 3 was set as zero point to compare relative stabilities of models 1–3.

### Competitive metal binding titrations

Equilibration dissociation constants of (AB)<sub>2</sub> with metal ions were obtained by competitive binding titrations using Fura-2, Mag-Fura-2 and Newport Green DCF<sup>25,31,66</sup>. Trace amounts of metal ions were removed from buffer solutions containing 20 mM MOPS and 150 mM NaCl (pH 7.4) using a Chelex 100 resin (Sigma-Aldrich). To a 1-ml buffer solution, 5 μM of metal-free dye and 5 μM of metal-free (AB)<sub>2</sub> were added, followed by the addition of aliquots of 2 mM MnCl<sub>2</sub>, CoCl<sub>2</sub>, NiCl<sub>2</sub>, CuCl<sub>2</sub> or ZnCl<sub>2</sub> stock solutions. Concentrations of all metal solutions were determined using a PAR assay<sup>67</sup>. Competitive binding of metal ions, dyes and (AB)<sub>2</sub> was monitored on an Agilent 8453 UV-vis spectrophotometer. Complete equilibration was achieved within 5 min following the addition of each metal aliquot. Changes in the absorbance at 335 nm (Fura-2), 324 nm (Mag-Fura-2) and 511 nm (Newport Green DCF) were used to measure the concentration of metal-bound indicators. Addition of metal aliquots was sequentially performed until the absorbances of the dyes were fully saturated (that is, no further changes were observed). All titrations were done in triplicate and the titration curves were fit to a one- or a two-site binding model for protein using Dynafit<sup>68</sup>, taking into account the total indicator concentration and the metal concentration at which the absorbance for the metal-bound indicator reached the saturation point.

### Electron paramagnetic resonance measurements

The electron paramagnetic resonance (EPR) sample (5 mM Cu<sup>2+</sup>-(AB)<sub>2</sub>) was prepared in a buffer solution containing 20 mM MOPS (pH 7.4) and 150 mM NaCl and loaded into a Wilmad Q-1.0 × 1.2 quartz capillary tube. The capillary tube was then placed in a 4-mm outer-diameter Wilmad quartz tube (707-SQ-250M). Continuous wave (CW) EPR data for Cu<sup>2+</sup>-(AB)<sub>2</sub> were collected at X-band frequency (9–9.8 GHz) on a Bruker EMXplus EPR spectrometer equipped with a high-sensitivity resonator (ER4119HS) at UCSD. The CW-EPR spectrum was collected at 298 K using Xenon software and fit using Easyspin software<sup>69</sup>.

## Calculation of metal-free and metal-bound (AB)<sub>2</sub> fractions under competitive conditions with Cu<sup>2+</sup>

Fractions of metal-free and metal-bound (AB)<sub>2</sub> species were calculated using  $K_d$  values of 1Cu<sup>2+</sup>-(AB)<sub>2</sub>, 2Ni<sup>2+</sup>-(AB)<sub>2</sub> and 2Co<sup>2+</sup>-(AB)<sub>2</sub> as a function of metal concentration. A mathematical model of the (AB)<sub>2</sub> fractions is available in the Supplementary Methods.

## Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

## Data availability

The principal data supporting the findings of this work are available within the figures and the Supplementary Information. Additional data that support the findings of this study are available from the corresponding author on request. Crystallographic data for protein structures (coordinates and structure factors) have been deposited into the RCSB PDB under the following accession codes: 7MK4, 7LRV, 7LV1, 7N4G, 7N4F, 7LRS, 7LRA, 7LRB and 7LRR. The model structure used for molecular replacement is available in the PDB (2BC5).

43. Arslan, E., Schulz, H., Zufferey, R., Künzler, P. & Thöny-Meyer, L. Overproduction of the *Bradyrhizobium japonicum* c-type cytochrome subunits of the *cbb<sub>3</sub>* oxidase in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **251**, 744–747 (1998).
44. Bailey, J. B., Subramanian, R. H., Churchfield, L. A. & Tezcan, F. A. in Peptide, Protein and Enzyme Design: Methods in Enzymology Vol. 580 (ed. Pecoraro, V. L.) 223–250 (Academic Press, 2016).
45. Martel, A., Liu, P., Weiss, T. M., Niebuhr, M. & Tsuruta, H. An integrated high-throughput data acquisition system for biological solution X-ray scattering studies. *J. Synchrotron Radiat.* **19**, 431–434 (2012).
46. Manalastas-Cantos, K. et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
47. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSTOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
48. Collaborative Computational Project. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
49. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
50. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
51. The PyMOL Molecular Graphics System v1.8 (Schrödinger, 2015).
52. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).
53. Manoil, C. & Beckwith, J. A genetic approach to analyzing membrane protein topology. *Science* **233**, 1403–1408 (1986).
54. Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K. & Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct. THEOCHEM* **461–462**, 1–21 (1999).
55. Vreven, T., Morokuma, K., Farkas, Ö., Schlegel, H. B. & Frisch, M. J. Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comput. Chem.* **24**, 760–769 (2003).
56. Tao, P. et al. Matrix metalloproteinase 2 inhibition: combined quantum mechanics and molecular mechanics studies of the inhibition mechanism of (4-phenoxyphenylsulfonyl)methylthiirane and its oxirane analogue. *Biochemistry* **48**, 9839–9847 (2009).
57. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
58. Lee, C., Yang, W. & Parr, R. G. Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
59. Hariharan, P. C. & Pople, J. A. The effect of d-functions on molecular orbital energies for hydrocarbons. *Chem. Phys. Lett.* **16**, 217–219 (1972).
60. Rassolov, V. A., Pople, J. A., Ratner, M. A. & Windus, T. L. 6-31G\* basis set for atoms K through Zn. *J. Chem. Phys.* **109**, 1223–1229 (1998).
61. Rassolov, V. A., Ratner, M. A., Pople, J. A., Redfern, P. C. & Curtiss, L. A. 6-31G\* basis set for third-row atoms. *J. Comput. Chem.* **22**, 976–984 (2001).
62. Freindorf, M., Shao, Y., Furlani, T. R. & Kong, J. Lennard–Jones parameters for the combined QM/MM method using the B3LYP/6-31G\*/AMBER potential. *J. Comput. Chem.* **26**, 1270–1278 (2005).
63. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).
64. Bakowies, D. & Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **100**, 10580–10594 (1996).
65. Weiner, S. J., Singh, U. C. & Kollman, P. A. Simulation of formamide hydrolysis by hydroxide ion in the gas phase and in aqueous solution. *J. Am. Chem. Soc.* **107**, 2219–2229 (1985).
66. Kakkis, A., Gagnon, D., Esselborn, J., Britt, R. D. & Tezcan, F. A. Metal-templated design of chemically switchable protein assemblies with high-affinity coordination sites. *Angew. Chem. Int. Ed. Engl.* **59**, 21940–21944 (2020).
67. Kocyla, A., Pomorski, A. & Kręzel, A. Molar absorption coefficients and stability constants of metal complexes of 4-(2-pyridylazo)resorcinol (PAR): revisiting common chelating probe for the study of metalloproteins. *J. Inorg. Biochem.* **152**, 82–92 (2015).
68. Kuzmić, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273 (1996).
69. Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42–55 (2006).
70. Smilgies, D.-M. & Folta-Stogniew, E. Molecular weight-gyration radius relation of globular proteins: a comparison of light scattering, small-angle X-ray scattering and structure-based data. *J. Appl. Crystallogr.* **48**, 1604–1606 (2015).

**Acknowledgements** We thank members of the Tezcan group, J. Rittle, H. Gray, S. Cohen and M. Green for discussions. We also thank A. Kakkis and C.-J. Yu for AUC and EPR measurements, respectively. This work was funded primarily by the National Institutes of Health (NIH; R01-GM138884) and by NASA (80NSSC18M0093; ENIGMA: Evolution of Nanomachines in Geospheres and Microbial Ancestors (NASA Astrobiology Institute Cycle 8)). Parts of this research were carried out at the Stanford Synchrotron Radiation Lightsource (supported by the DOE, Office of Basic Energy Sciences contract DE-AC02-76SF00515 and NIH P30-GM133894) and the Advanced Light Source (supported by the DOE, Office of Basic Energy Sciences contract DE-AC02-05CH11231 and NIH P30-GM124169-01). EPR data was acquired on an instrument obtained through a Major Research Instrumentation fund supported by the National Science Foundation (NSF; CHE-2019066).

**Author contributions** T.S.C. conceived the project, designed and performed all experiments, calculations and analyses and co-wrote the paper. F.A.T. conceived and directed the project and co-wrote the manuscript.

**Competing interests** The authors declare no competing interests.

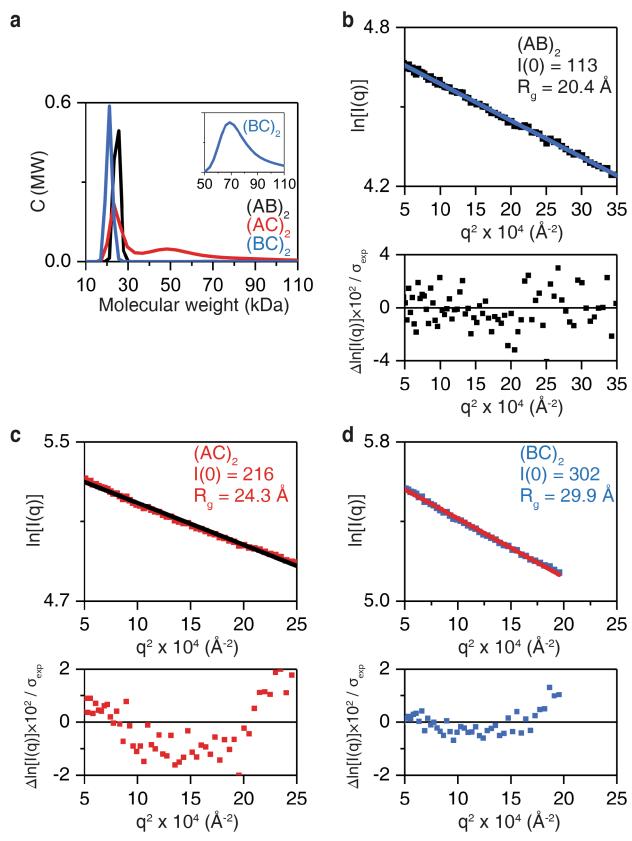
## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41586-022-04469-8>.

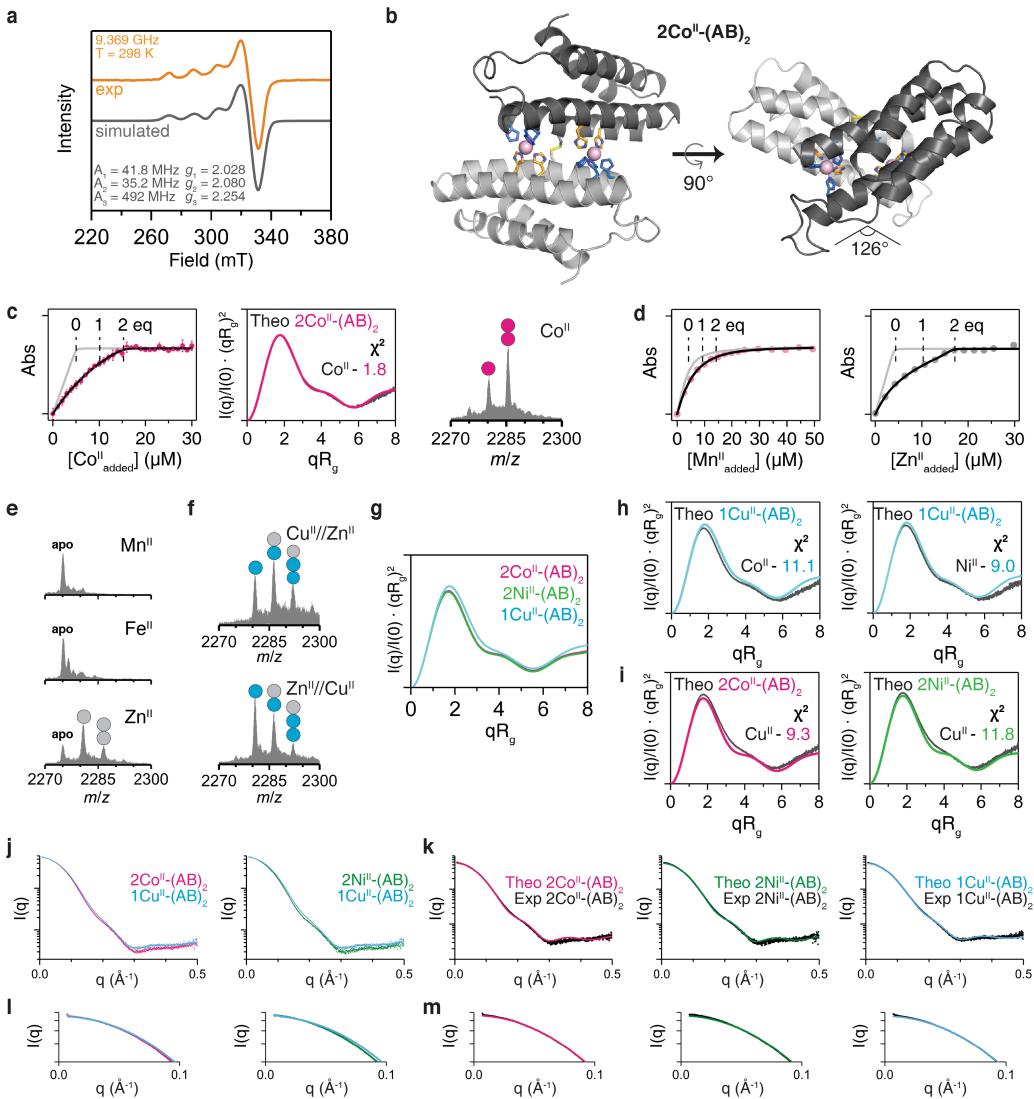
**Correspondence and requests for materials** should be addressed to F. Akif Tezcan.

**Peer review information** *Nature* thanks Jan Lipfert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>.

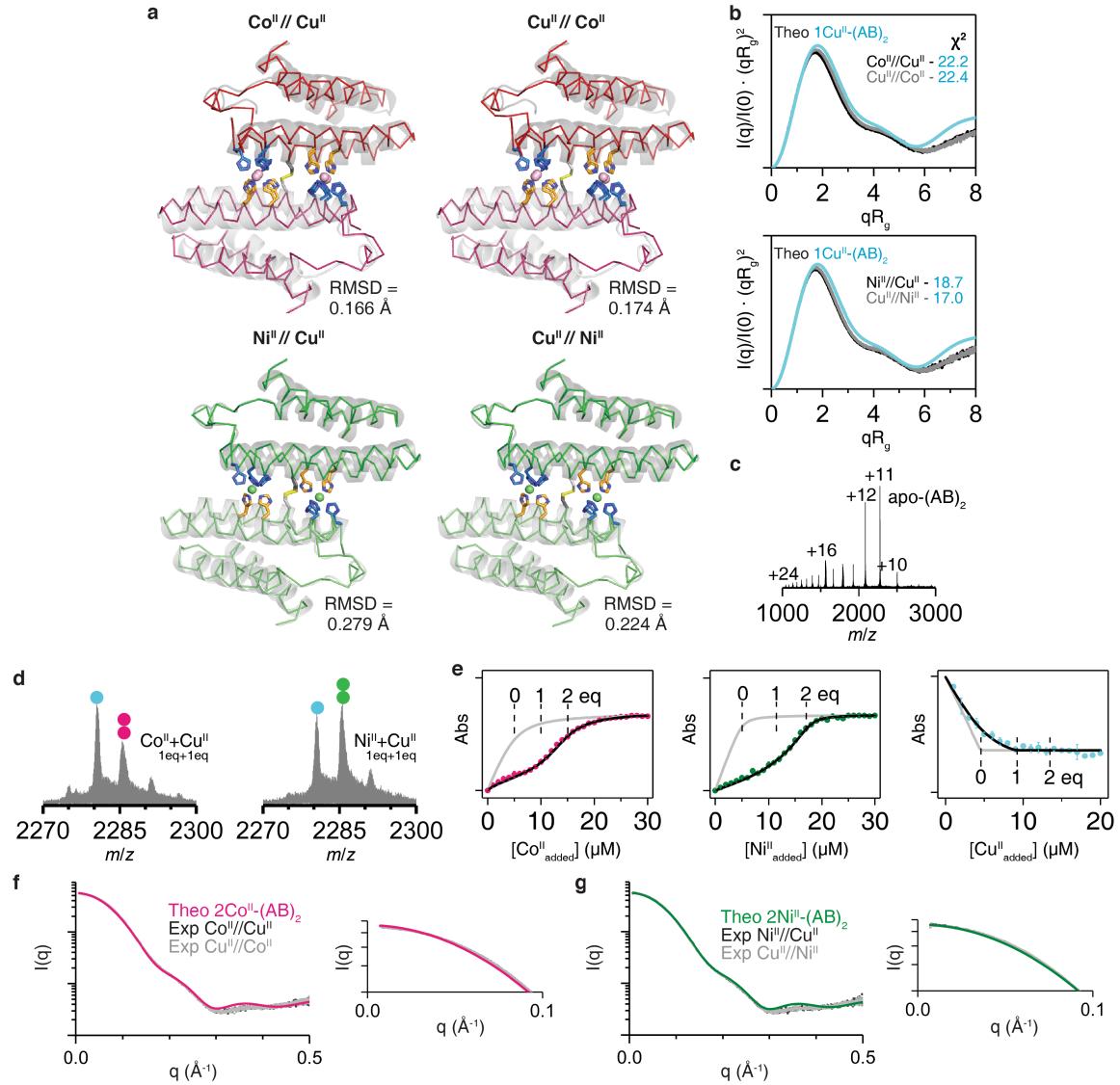


**Extended Data Fig. 1 | Analysis of the oligomerization states of  $(AB)_2$ ,  $(AC)_2$  and  $(BC)_2$  variants.** **a**, Sedimentation velocity/AUC analysis of the oligomerization state of  $(AB)_2$  (black),  $(AC)_2$  (red), and  $(BC)_2$  (blue) at 25  $\mu\text{M}$  dimer concentration. The inset shows a higher-order oligomer population of  $(BC)_2$ . **b–d**, Guinier analysis of **(b)** AB (black), **(c)** AC (red), and **(d)** BC (blue) constructs measured by solution small-angle X-ray scattering (SAXS) at 0.8 mM dimer concentration. Larger  $I(0)$  values in  $(AC)_2$  and  $(BC)_2$  indicate higher oligomeric states compared to  $(AB)_2$ . Radius of gyration ( $R_g$ ) value of  $(AB)_2$  corresponds to approximately 25 kDa in  $M_w - R_g$  relationship of globular proteins<sup>70</sup>, indicating a dimeric state for  $(AB)_2$ .



**Extended Data Fig. 2 | Structural characterization and metal-binding analysis of M<sup>II</sup>-(AB)<sub>2</sub> complexes.** **a**, X-band EPR spectrum of Cu<sup>II</sup>-(AB)<sub>2</sub> (orange) and simulated pattern (black line), along with the fit parameters consistent with a rhombic coordination geometry. Conditions: 2.5 mM Cu<sup>II</sup>, 20 mM MOPS (pH 7.4) and 150 mM NaCl, 298 K. **b**, Crystal structure of 2Co<sup>II</sup>-(AB)<sub>2</sub>. Co<sup>II</sup> ions are represented as magenta spheres. **c**, Solution characterization of 2Co<sup>II</sup>-(AB)<sub>2</sub> complex by competitive Fura-2 titration (left), SAXS (middle), and ESI-MS (right). Circles (magenta) in ESI-MS spectrum represent the number of Co<sup>II</sup> ions bound to (AB)<sub>2</sub>. Experimental data points and error bars in the Fura-2 titration are presented as mean and standard deviation of three independent measurements. **d**, Competitive metal-binding titrations (AB)<sub>2</sub> for Mn<sup>II</sup> (pink) and Zn<sup>II</sup> (grey) binding. Mag-Fura-2 and Fura-2 were used for Mn<sup>II</sup> and Zn<sup>II</sup> titrations, respectively. **e**, Investigation of M<sup>II</sup>-(AB)<sub>2</sub> complexation using ESI-MS:

Mn<sup>II</sup> (top), Fe<sup>II</sup> (middle), and Zn<sup>II</sup> (bottom). Mn<sup>II</sup>- and Fe<sup>II</sup>-(AB)<sub>2</sub> complexes were not observed owing to the low metal-binding affinities. **f**, ESI-MS spectra of (AB)<sub>2</sub> collected under Cu<sup>II</sup>//Zn<sup>II</sup> and Zn<sup>II</sup>//Cu<sup>II</sup> competition conditions. Cu<sup>II</sup> outcompetes Zn<sup>II</sup> or forms heterometallic complexes of (AB)<sub>2</sub>. **g**, Theoretical SAXS scattering profiles of M<sup>II</sup>-(AB)<sub>2</sub> generated by CRYSTAL. **h**, **i**, Experimental SAXS profiles of **(h)** 2Co<sup>II</sup>- and 2Ni<sup>II</sup>-(AB)<sub>2</sub> complexes compared with theoretical scattering profile of 1Cu<sup>II</sup>-(AB)<sub>2</sub> (cyan), and **(i)** 1Cu<sup>II</sup>-(AB)<sub>2</sub> complex compared with theoretical scattering profiles of 2Co<sup>II</sup>- (magenta) and 2Ni<sup>II</sup>-(AB)<sub>2</sub> (green). **j**, **k**, Log-scale plots of Fig. 1e–f to compare **(j)** experimental scattering profiles between 2Co<sup>II</sup>/2Ni<sup>II</sup>-(AB)<sub>2</sub> and 1Cu<sup>II</sup>-(AB)<sub>2</sub>, and **(k)** experimental scattering profiles of 2Co<sup>II</sup>- (left), 2Ni<sup>II</sup>- (middle), and 1Cu<sup>II</sup>-(AB)<sub>2</sub> (right) with theoretical scattering profiles. **l**, **m**, Expanded low q-ranges of the scattering plots in **(l)** (panel **j**) and **(k)**.

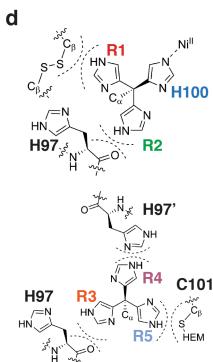
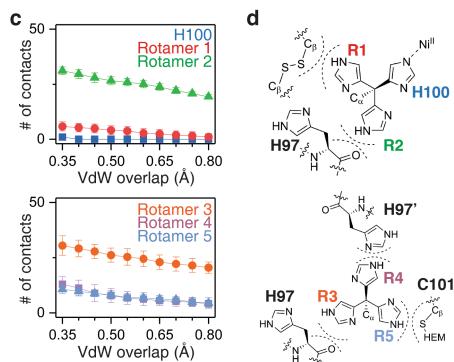
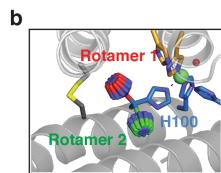
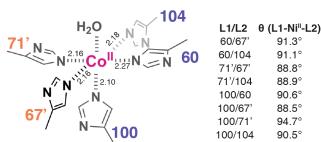
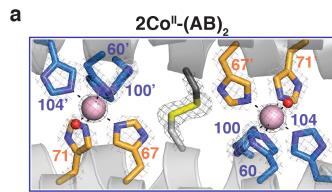


**Extended Data Fig. 3 | Structural characterization and metal-binding analysis of M<sup>II</sup>-(AB)<sub>2</sub> complexes in competitive conditions with Cu<sup>II</sup>.**

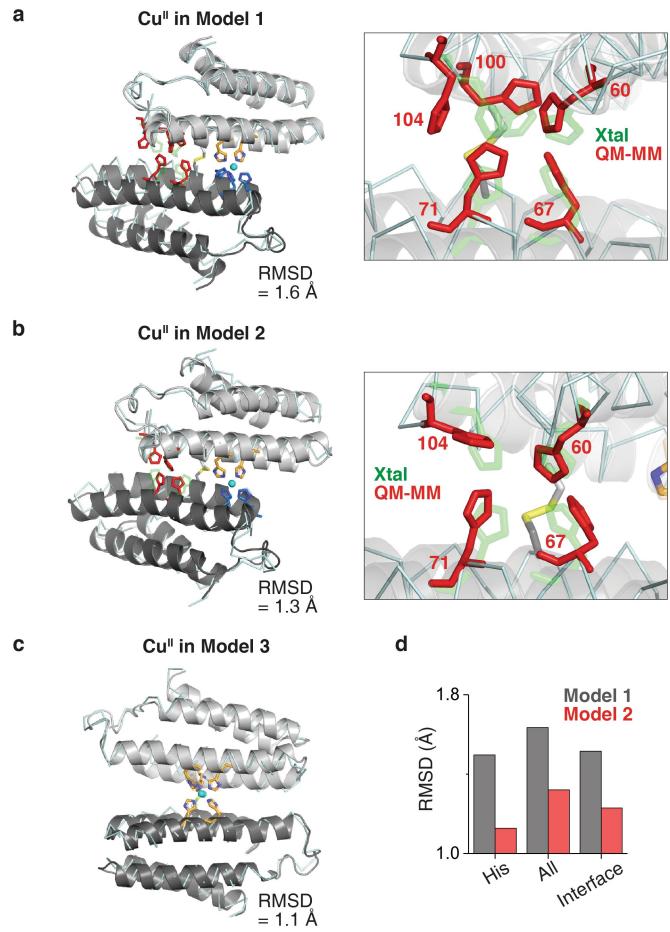
**a**, Crystal structures (ribbon models) of (AB)<sub>2</sub> formed under Co<sup>II</sup>//Cu<sup>II</sup> (PDB:7LRB), Cu<sup>II</sup>//Co<sup>II</sup> (PDB:7LRR), Ni<sup>II</sup>//Cu<sup>II</sup> (PDB:7LR5), and Cu<sup>II</sup>//Ni<sup>II</sup> (PDB:7LRA) conditions. Root mean square deviation (RMSD) values were determined in comparison with the crystal structures of 2Co<sup>II</sup>-(AB)<sub>2</sub> and 2Ni<sup>II</sup>-(AB)<sub>2</sub> (grey cartoon models). **b**, Experimental SAXS profile of (AB)<sub>2</sub> (0.8 mM) in Co<sup>II</sup>//Cu<sup>II</sup>, Cu<sup>II</sup>//Co<sup>II</sup>, Ni<sup>II</sup>//Cu<sup>II</sup>, and Cu<sup>II</sup>//Ni<sup>II</sup> conditions ([M<sup>II</sup>] = 1.6 mM) compared with the theoretical SAXS profile of 1Cu<sup>II</sup>-(AB)<sub>2</sub>. **c**, ESI-MS spectrum of (AB)<sub>2</sub> (5 μM) without metal ions. **d**, ESI-MS spectrum of (AB)<sub>2</sub> (5 μM) with sub-stoichiometric amounts of Cu<sup>II</sup> (5 μM) and Co<sup>II</sup> or Ni<sup>II</sup> (5 μM). Circles in ESI-MS spectra

represent the number of Co<sup>II</sup> (magenta), Ni<sup>II</sup> (green), and Cu<sup>II</sup> (cyan) ions bound to (AB)<sub>2</sub>. The number of circles indicates the equivalents of bound metal ions. **e**, Competitive metal-binding titration of (AB)<sub>2</sub> with Co<sup>II</sup> (magenta), Ni<sup>II</sup> (green), and Cu<sup>II</sup> (cyan) in 20 mM NH<sub>4</sub>HCO<sub>3</sub> (pH 7.8). Mag-Fura-2 (5 μM) was used for competitive Co<sup>II</sup> and Ni<sup>II</sup> titration, and Newport green DCF (5 μM) was used for competitive Cu<sup>II</sup> titration. Experimental data points and error bars are presented as mean and standard deviation of three independent measurements. **f**, **g**, Log-scale plots (left) of Fig. 2c to compare experimental scattering profiles of (f) Co<sup>II</sup>//Cu<sup>II</sup> and Cu<sup>II</sup>//Co<sup>II</sup>, and (g) Ni<sup>II</sup>//Cu<sup>II</sup> and Cu<sup>II</sup>//Ni<sup>II</sup> with theoretical scattering profiles. Right panels present expanded low q-ranges of the scattering plots (left).

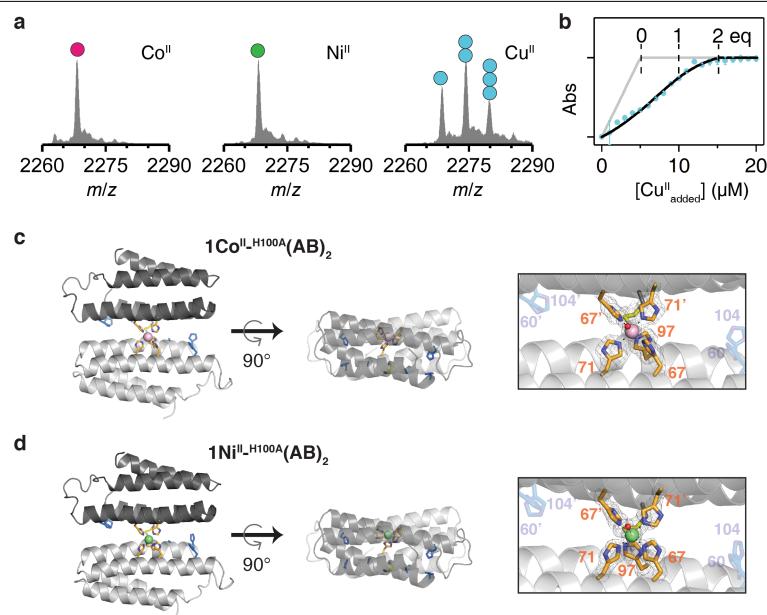
# Article



**Extended Data Fig. 4 | Peripheral 5His coordination site of  $2\text{Co}^{\text{II}}\text{-}(\text{AB})_2$ , and rotamer analysis of the H100 side chain.** **a**,  $\text{Co}^{\text{II}}$ -coordination in the dimer interface of  $(\text{AB})_2$  with the  $2\text{mF}_\alpha\text{-DF}_\epsilon$  electron density map (grey mesh) contoured at  $5.0\sigma$  (metal) and  $1.5\sigma$  (ligand). Coordination distances and angles between  $\text{Co}^{\text{II}}$  and ligands are shown on the right. **b**, Possible rotamer orientations of H100 residue in the crystal structure of  $2\text{Ni}^{\text{II}}\text{-}(\text{AB})_2$ . Rotamers were generated with the combination of  $\chi_1$  (torsion angle of  $\text{C}_\alpha\text{-C}_\beta = 60\text{--}300^\circ$ ,  $60^\circ$  interval) and  $\chi_2$  (torsion angle of  $\text{C}_\beta\text{-C}_\gamma = 0\text{--}330^\circ$ ,  $30^\circ$  interval). **c**, Averaged number of atomic contacts of H100 rotamers as a function of van der waals (VdW) overlap. Error bars reflect the standard deviations of the atomic contacts in rotamers with fixed  $\chi_1$  and variable  $\chi_2$ . Compared to the original conformation of H100, all possible rotamers show significant clashes with proximal residues. **d**, Location of proximal residues interacting with H100 rotamers.

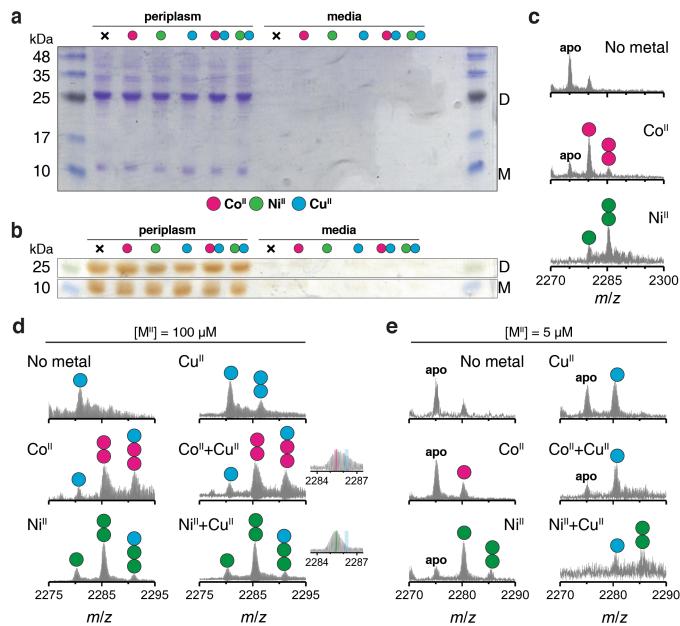


**Extended Data Fig. 5 | QM/MM calculations of models 1–3.** **a–c**, QM/MM-optimized structures of **(a)** Model 1: Cu<sup>II</sup> in the peripheral site of (AB)<sub>2</sub>, **(b)** Model 2: Cu<sup>II</sup> in the peripheral site of <sup>H100A</sup>(AB)<sub>2</sub>, and **(c)** Model 3: Cu<sup>II</sup> in the central site of (AB)<sub>2</sub>. Conformations of the opposite peripheral binding site (red) from the modelled Cu<sup>II</sup> coordination site are shown in the panel on the right in **a** and **b**. **d**, RMSDs of C<sub>α</sub> positions between QM/MM-optimized models and the crystal structure (His: C<sub>α</sub>'s of 5His or 4His residues in the opposite peripheral site, All: all C<sub>α</sub>'s in (AB)<sub>2</sub> or <sup>H100A</sup>(AB)<sub>2</sub>, and Interface: C<sub>α</sub>'s of residues in the dimer interface).



**Extended Data Fig. 6 | Characterization of metal-bound  ${}^{\text{H}100\text{A}}(\text{AB})_2$  complexes.** **a**, ESI-MS spectra of  ${}^{\text{H}100\text{A}}(\text{AB})_2$  ( $5 \mu\text{M}$ ) complexed with metal ions. Circles in ESI-MS spectra represent the number of  $\text{Co}^{\text{II}}$  (magenta),  $\text{Ni}^{\text{II}}$  (green), and  $\text{Cu}^{\text{II}}$  (cyan) ions bound to  ${}^{\text{H}100\text{A}}(\text{AB})_2$ . The number of circles indicates the equivalents of bound metal ions. Metal concentrations were  $10 \mu\text{M}$ . **b**, Fura-2 competitive metal titration assay of  ${}^{\text{H}100\text{A}}(\text{AB})_2$  with  $\text{Cu}^{\text{II}}$ . Changes in Fura-2 absorbance at  $335 \text{ nm}$  (cyan) are plotted with theoretical metal-binding

isotherms in the absence (grey) and the presence (black) of  $(\text{AB})_2$ . Experimental data points and error bars in the Fura-2 titration are presented as mean and standard deviation of three independent measurements. **c, d**, Crystal structure of (c)  $\text{Co}^{\text{II}}\text{-}{}^{\text{H}100\text{A}}(\text{AB})_2$  (PDB:7N4G) and (d)  $\text{Ni}^{\text{II}}\text{-}{}^{\text{H}100\text{A}}(\text{AB})_2$  (PDB:7N4F).  $\text{Co}^{\text{II}}$  and  $\text{Ni}^{\text{II}}$  ions are represented as magenta and green spheres. Atomic details of each metal coordination site are shown in the right panels, with the  $2\text{mF}_o\text{-DF}_c$  electron density map (grey mesh) contoured at  $5.0\sigma$  (metal) and  $1.5\sigma$  (ligand).



**Extended Data Fig. 7 | Characterization of periplasmic extracts using SDS-PAGE and ESI-MS.** **a, b**, SDS-PAGE of periplasmic extracts and medium stained with (a) Coomassie blue for all proteinaceous contents and (b) o-dianisidine for haem-proteins. Regardless of supplemented metal ions, no significant amount of  $(AB)_2$  dimer and  $(AB)$  monomer was observed in the medium. Uncropped gel images are shown in Supplementary Fig. 2. Reproducibility of SDS-PAGE was tested with two independently extracted sample sets. **c**, ESI-MS spectra of  $(AB)_2$  extracted from cells grown without metal supplement (top) and with Co<sup>II</sup> (middle) or Ni<sup>II</sup> (bottom) supplement (20  $\mu$ M) in LB medium. **d, e**, ESI-MS spectra of  $(AB)_2$  extracted from cells grown with (d) 100  $\mu$ M and (e) 5  $\mu$ M metal ions in LB medium. Inlet spectra represent expanded  $m/z$  ranges for  $2M^{II}-(AB)_2$  complexes with magenta, green, and cyan lines corresponding to theoretical  $m/z$  values of  $2Co^{II}$ ,  $2Ni^{II}$ , and  $2Cu^{II}-(AB)_2$  complexes. Supplemented metal ions are represented as magenta (Co<sup>II</sup>), green (Ni<sup>II</sup>), and cyan (Cu<sup>II</sup>) circles. The number of circles indicates the equivalents of metal ions bound to  $(AB)_2$ . D and M in gel pictures refer to the protein dimer and monomer, respectively.

# Article

**Extended Data Table 1 | Equilibrium dissociation constants ( $K_d$ ) of M<sup>II</sup>-(AB)<sub>2</sub> complexes**

		(AB) <sub>2</sub>	(AB) <sub>2</sub> (20 mM NH <sub>4</sub> HCO <sub>3</sub> )
<b>Mn<sup>II</sup></b>	$K_{d1}$	2.7 ( $\pm 0.4$ ) $\times 10^{-6}$	
	$K_{d2}$	4.3 ( $\pm 1.1$ ) $\times 10^{-6}$	
<b>Co<sup>II</sup></b>	$K_{d1}$	1.1 ( $\pm 0.2$ ) $\times 10^{-8}$	5.2 ( $\pm 1.8$ ) $\times 10^{-8}$
	$K_{d2}$	4.1 ( $\pm 0.7$ ) $\times 10^{-8}$	1.1 ( $\pm 0.3$ ) $\times 10^{-7}$
<b>Ni<sup>II</sup></b>	$K_{d1}$	4.5 ( $\pm 1.1$ ) $\times 10^{-9}$	3.0 ( $\pm 0.7$ ) $\times 10^{-8}$
	$K_{d2}$	7.6 ( $\pm 1.0$ ) $\times 10^{-9}$	6.1 ( $\pm 0.7$ ) $\times 10^{-8}$
<b>Cu<sup>II</sup></b>	$K_{d1}$	2.9 ( $\pm 0.5$ ) $\times 10^{-13}$	4.3 ( $\pm 1.4$ ) $\times 10^{-9}$
	$K_{d2}$	-	-
<b>Zn<sup>II</sup></b>	$K_{d1}$	1.4 ( $\pm 0.3$ ) $\times 10^{-8}$	
	$K_{d2}$	1.3 ( $\pm 0.4$ ) $\times 10^{-8}$	

Metal-competition assays were performed using Fura-2 in Chelex100-treated buffer solution (pH 7.4) containing 20 mM MOPS and 150 mM NaCl.  $K_d$  values in 20 mM NH<sub>4</sub>HCO<sub>3</sub> were determined using Mag-Fura-2 (Co<sup>II</sup> and Ni<sup>II</sup>) and Newport Green DCF (Cu<sup>II</sup>) as metal-binding competitor. Errors are derived from non-linear least square fitting analysed in Dynafit.

**Extended Data Table 2 | X-ray data collection and refinement statistics for the crystals obtained under non-competitive conditions**

PDB ID	2Co <sup>II</sup> -(AB) <sub>2</sub> 7MK4	2Ni <sup>II</sup> -(AB) <sub>2</sub> 7LRV	1Cu <sup>II</sup> -(AB) <sub>2</sub> 7LV1	1Co <sup>II</sup> -H100A(AB) <sub>2</sub> 7N4G	1Ni <sup>II</sup> -H100A(AB) <sub>2</sub> 7N4F
<b>Data collection</b>					
Resol. range (Å) (1.315 - 1.269)	34.42 - 1.269 (1.315 - 1.269)	41.61 - 1.4 (1.45 - 1.4)	37.88 - 1.91 (1.978 - 1.91)	34.6 - 1.93 (1.999 - 1.93)	35.03 - 1.8 (1.864 - 1.8)
Space group	P 21 21 21	P 21 21 21	P 1 21 1	P 1 21 1	P 1 21 1
Unit cell (Å, °)	54.53 61.08 64.57 90 90 90	54.91 60.72 63.77 90 90 90	34.17 83.21 38.57 90 100.83 90	35.21 78.27 39.90 100.68 90	34.47 84.65 39.14 90 100.51 90
Total reflections	416194 (26635)	1371309 (118524)	131503 (8872)	217256 (21702)	242223 (24300)
Unique reflections	57551 (5614)	42610 (4185)	16289 (1530)	15651 (1535)	20432 (1985)
Multiplicity	7.2 (4.7)	32.2 (28.3)	8.07 (5.8)	13.9 (14.0)	11.9 (11.8)
Completeness (%)	99.79 (98.86)	99.80 (99.12)	98.82 (93.62)	98.89 (98.65)	96.31 (96.60)
Mean I/sigma(I)	10.0 (1.5)	152.76 (9.62)	20.9 (5.8)	159.03 (6.55)	271.29 (19.40)
Wilson B-factor	14	16.52	15.15	36.53	25.88
R-merge	0.067 (0.847)	0.744 (0.9591)	0.063 (0.214)	0.5892 (0.8141)	0.5845 (0.5606)
R-meas	0.072 (0.953)	0.7563 (0.9731)	0.067 (0.237)	0.6126 (0.8456)	0.6132 (0.5889)
R-pim	0.024 (0.429)	0.1333 (0.1631)	0.022 (0.099)	0.1643 (0.2242)	0.1812 (0.1773)
CC1/2	0.999 (0.765)	0.738 (0.804)	0.999 (0.978)	0.82 (0.83)	0.749 (0.817)
<b>Refinement</b>					
Reflections used in refinement	57550 (5612)	42605 (4183)	16289 (1511)	15519 (1533)	19738 (1986)
Reflections used for R-free	2002 (191)	1995 (193)	1625 (150)	1549 (153)	1974 (199)
R-work	0.1960 (0.3160)	0.1901 (0.2097)	0.1908 (0.2724)	0.2640 (0.3957)	0.2078 (0.2966)
R-free	0.2135 (0.3582)	0.2161 (0.2557)	0.2371 (0.2766)	0.3245 (0.4594)	0.2537 (0.3489)
Number of non-hydrogen atoms	2123	1975	1927	1757	1801
macromolecules	1671	1636	1670	1651	1640
ligands	90	88	91	87	87
solvent	362	251	166	19	74
Protein residues	214	214	214	214	214
RMS(bonds)	0.006	0.006	0.008	0.009	0.008
RMS(angles)	0.81	0.78	0.84	1.19	0.93
Ramachandran favored (%)	100	99.04	99.52	96.63	99.52
Ramachandran allowed (%)	0	0.96	0.48	3.37	0.48
Ramachandran outliers (%)	0	0	0	0	0
Rotamer outliers (%)	0	0	0	0	0.61
Clashscore	1.76	1.8	3.79	12.11	4.48
Average B-factor	19.01	23.24	18.1	49.13	32.97
macromolecules	17.84	22.77	17.93	49.45	33.31
ligands	15.22	21.96	13.72	46.2	26.74
solvent	25.36	26.73	22.24	34.22	32.78

Numbers in parentheses correspond to statistics of the highest resolution shell.

# Article

**Extended Data Table 3 | X-ray data collection and refinement statistics for the crystals obtained under competitive conditions with Cu<sup>II</sup>**

	2Ni <sup>II</sup> -(AB) <sub>2</sub> (Ni <sup>II</sup> //Cu <sup>II</sup> )	2Ni <sup>II</sup> -(AB) <sub>2</sub> (Cu <sup>II</sup> //Ni <sup>II</sup> )	2Co <sup>II</sup> -(AB) <sub>2</sub> (Co <sup>II</sup> //Cu <sup>II</sup> )	2Co <sup>II</sup> -(AB) <sub>2</sub> (Cu <sup>II</sup> //Co <sup>II</sup> )
PDB ID	7LR5	7LRA	7LRB	7LRR
<b>Data collection</b>				
Resol. range (Å)	41.32 - 1.7 (1.761 - 1.7)	41.65 - 1.7 (1.761 - 1.7)	41.63 - 1.78 (1.844 - 1.78)	44.27 - 1.89 (1.958 - 1.89)
Space group	P 21 21 21			
Unit cell (Å, °)	54.63 60.92 63.16 90 90 90	54.76 61 64.16 90 90 90	54.63 61.05 64.30 90 90 90	54.51 61.15 64.17 90 90 90
Total reflections	919338 (85881)	686593 (63701)	541511 (41580)	531273 (31757)
Unique reflections	23765 (2334)	24213 (2371)	20861 (1911)	17578 (1611)
Multiplicity	38.7 (36.7)	28.4 (26.8)	26.0 (21.6)	30.2 (19.4)
Completeness (%)	99.50 (99.62)	99.54 (99.37)	97.65 (91.70)	98.54 (92.48)
Mean I/sigma(I)	35.10 (6.50)	49.46 (8.63)	31.17 (6.16)	14.77 (3.30)
Wilson B-factor	17.96	17.98	17.17	16.99
R-merge	0.7238 (0.9307)	0.7754 (0.8503)	0.7359 (0.7766)	0.8529 (0.9726)
R-meas	0.7335 (0.9442)	0.79 (0.8665)	0.7509 (0.7961)	0.8678 (0.9993)
R-pim	0.1161 (0.156)	0.1475 (0.1641)	0.1447 (0.1692)	0.156 (0.223)
CC1/2	0.878 (0.727)	0.73 (0.771)	0.676 (0.788)	0.883 (0.417)
<b>Refinement</b>				
Reflections used in refinement	23704 (2333)	24165 (2371)	20721 (1911)	17488 (1611)
Reflections used for R-free	2001 (196)	2010 (198)	1995 (180)	1762 (158)
R-work	0.1893 (0.2192)	0.1802 (0.2215)	0.1774 (0.2232)	0.1902 (0.2544)
R-free	0.2296 (0.2790)	0.2141 (0.2508)	0.2112 (0.2791)	0.2294 (0.3165)
Number of non-hydrogen atoms	1899	1968	1990	1940
macromolecules	1637	1635	1680	1667
ligands	88	88	90	90
solvent	174	245	220	183
Protein residues	214	214	214	214
RMS(bonds)	0.007	0.006	0.006	0.007
RMS(angles)	0.8	0.78	0.86	0.82
Ramachandran favored (%)	99.52	99.52	100	100
Ramachandran allowed (%)	0.48	0.48	0	0
Ramachandran outliers (%)	0	0	0	0
Rotamer outliers (%)	0	0	0.58	0
Clashscore	3	4.51	2.91	2.64
Average B-factor	20.56	21.82	19.86	19.43
macromolecules	20.34	21.23	19.41	19.26
ligands	19.44	20.83	17.83	17.48
solvent	23.22	26.13	24.08	21.93

Numbers in parentheses correspond to statistics of the highest resolution shell

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
  - Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted
  - Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

#### Data collection

X-ray crystallography: Blu-Ice 5 at SSRL and ALS APEX3 v2016 at UCSD  
Solution small-angle X-ray scattering: Blu-Ice 5 in SSRL 4-2 beamline  
Electrospray ionization mass spectrometry: Xcalibur 4.3.73.11  
UV/Vis: 8453 Agilent UV-Visible System  
AUC: Proteomelab  
EPR: Xenon for Bruker EMXplus EPR spectrometer

#### Data analysis

X-ray crystallography: imosflm 7.2.2 (data integration), Scala (data reduction/scaling), Phaser-MR (molecular replacement), Phenix 1.15.2-3472 (model building/refinement), Coot v.0.8.9.2 (refinement), and PyMOL version 1.3 (molecular graphics)  
Solution small-angle X-ray scattering: SAXSPipe at SSRL and ATSAS package 2.7.2  
Electrospray ionization mass spectrometry: Xcalibur 4.3.73.11  
UV/Vis:Dynafit 4.08.161  
AUC: SEDFIT software 16-1c  
QM-MM: Gaussian09 (QM/MM - B3LYP/Amber96) and TAO package  
EPR: Easyspin 5.2.33

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

## Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The principal data supporting the findings of this work are available within the figures and the Supplementary Information. Additional data that support the findings of this study are available from the corresponding author on request. Crystallographic data for protein structures (coordinates and structure factors) have been deposited into the RCSB data bank under the following accession codes: 7MK4, 7LRV, 7LV1, 7N4G, 7N4F, 7LR5, 7LRA, 7LRB and 7LRR. The model structure used for molecular replacement of processed X-ray diffraction data is also available in the RCSB data bank (code: 2BC5).

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences       Behavioural & social sciences       Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Various sample sizes for different experiments were selected. Metal-binding properties of artificially designed proteins in vitro and in bacteria cells were reproducible with the sample sizes.
Data exclusions	No data was excluded.
Replication	All experiments were successfully replicated within two weeks or sequential synchrotron beamtimes (The beamtimes were allocated with the interval of 3–6 months).
Randomization	Qualitative analysis of artificially designed proteins (structure/stoichiometry) performed in this study is not relevant to randomization.
Blinding	Because structural/analytical characterizations of artificially designed proteins were performed, blinding is not relevant to this study. Instead, suitable control groups were included in all experimental results.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies
<input checked="" type="checkbox"/>	Eukaryotic cell lines
<input checked="" type="checkbox"/>	Palaeontology and archaeology
<input checked="" type="checkbox"/>	Animals and other organisms
<input checked="" type="checkbox"/>	Human research participants
<input checked="" type="checkbox"/>	Clinical data
<input checked="" type="checkbox"/>	Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	MRI-based neuroimaging